

Universidade do Minho
Escola de Economia e Gestão

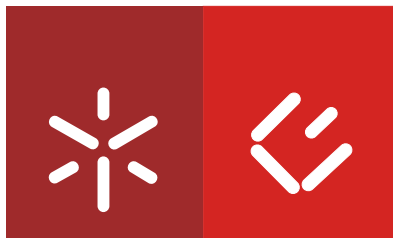
Joana Patrícia Martins Barbosa

Impacto do capital humano na sobrevivência das empresas

Joana Patrícia Martins Barbosa Impacto do capital humano na sobrevivência das empresas

UMinho | 2012

Outubro de 2012



Universidade do Minho

Escola de Economia e Gestão

Joana Patrícia Martins Barbosa

Impacto do capital humano na sobrevivência das empresas

Dissertação de Mestrado
Mestrado em Economia Industrial e da Empresa

Trabalho realizado sob a orientação da
Professora Ana Paula Rodrigues Pereira Faria
e do
Professor Miguel Ângelo dos Reis Portela

Outubro de 2012

É AUTORIZADA A REPRODUÇÃO PARCIAL DESTA DISSERTAÇÃO APENAS PARA EFEITOS DE INVESTIGAÇÃO, MEDIANTE DECLARAÇÃO ESCRITA DO INTERESSADO, QUE A TAL SE COMPROMETE;

Universidade do Minho, ____/____/____

Assinatura: _____

AGRADECIMENTOS

A presente dissertação de mestrado marca o final de uma etapa relevante no meu percurso académico não podendo esta ser encerrada sem o contributo de várias pessoas.

Aos professores doutores Ana Paula Rodrigues Pereira Faria e Miguel Ângelo Reis Portela, meus orientadores, agradeço a humildade, a dedicação, o profissionalismo e a partilha de conhecimentos valiosos.

Agradeço à Universidade do Minho, em particular à Escola de Economia e Gestão pela oportunidade e pela disponibilização do Laboratório de Informática – Sala dos Quadros de Pessoal.

Aos meus pais pelo investimento e confiança.

Às colegas com as quais partilhei o espaço de trabalho pela troca de conhecimentos.

A todos aqueles cujo nome não foi mencionado mas que me presentearam com o seu apoio.

Um sincero agradecimento a todos.

RESUMO

Vários autores, entre os quais Mata e Portugal (1994 e 2002), têm-se debruçado sobre a análise dos determinantes da sobrevivência das empresas sendo possíveis várias abordagens: determinantes relacionados com os trabalhadores, com as empresas ou com o ambiente no qual estas se inserem. À semelhança de alguns trabalhos analisados, como Bates (1990), a presente dissertação investiga o impacto do capital humano sobre essa mesma sobrevivência com recurso a modelos não paramétricos, semi-paramétricos e paramétricos aplicados à base de dados dos Quadros de Pessoal. O contributo distinto prende-se com a utilização de um índice de medição do capital humano que não foi utilizado por nenhum dos autores analisados.

Este tema é pertinente para empreendedores que ponderam criar uma empresa de sucesso e que questionam a importância do capital humano a contratar, assim como do ponto de vista da política económica dado que a entrada de novas empresas é um mecanismo importante na criação de emprego e crescimento económico.

Os resultados obtidos mostram que, no geral, quanto maior o ICH dos indivíduos, menor o rácio de risco das empresas. Assim sendo, como principal conclusão desta dissertação podemos afirmar que o capital humano tem um impacto positivo sobre a sobrevivência das empresas.

Palavras-chave: Índice de Capital Humano, Sobrevivência das Empresas, Modelo de duração

ABSTRACT

Several authors, including Mata and Portugal (1994 and 2002), have focused on the analysis of firms' survival determinants with several possible approaches: determinants related to workers, to firms or to the environment to which they belong. Like some studies analyzed, as Bates (1990), this dissertation investigates the human capital impact on the same survival using non-parametric, semi-parametric and parametric models applied to *Quadros de Pessoal* database. The distinct contribute relates to the use of an index measuring human capital that wasn't used by any of the authors analyzed.

This topic is relevant for entrepreneurs who are considering starting a successful business and who question the importance of human capital to hire, as well as from the standpoint of economic policy since the entry of new firms is an important mechanism in jobs creation and growth economic.

The results obtained show that, in general, as higher the human capital, lower the firms' hazard ratio. Therefore, the main conclusion of this dissertation is that human capital has a positive impact on firms' survival.

Keywords: Human Capital Index, Firms' Survival, Duration Model

ÍNDICE GERAL

CAPÍTULO I - INTRODUÇÃO	1
1.1. Contextualização e motivação	1
1.2. Objetivos e questões de investigação	2
1.3. Estrutura da dissertação	2
CAPÍTULO II - ENQUADRAMENTO TEÓRICO	3
2.2. Capital humano e sobrevivência das empresas	3
2.3. Determinantes da sobrevivência específicos à empresa	5
2.4. Determinantes setoriais	7
CAPÍTULO III - DADOS, VARIÁVEIS E MODELO EMPÍRICO	11
3.2. Fontes e definição das amostras	11
3.3. Modelo empírico e variáveis empíricas	13
CAPÍTULO IV - RESULTADOS	21
4.2. Modelo Econométrico	21
4.2.1. Análise não paramétrica	26
4.2.2. Análise semi-paramétrica	27
4.2.3. Análise paramétrica	29
4.3. Resultados	32
4.3.1. Modelos não paramétricos	32
4.3.2. Modelos semi-paramétricos e paramétricos	38
CAPÍTULO V - CONCLUSÕES	45
5.1. Síntese	45
5.2. Principais conclusões e recomendações	45
5.3. Limitações e pistas para investigação futura	47
BIBLIOGRAFIA	49
ANEXOS	55

ABREVIATURAS E SIGLAS

AIC – Akaike Information Criterion

BIC – Bayesian Information Criterion

I&D – Investigação de Desenvolvimento

ICH – Índice de Capital Humano

IHH - Índice de Herfindahl-Hirschman

INE – Instituto Nacional de Estatística

MEE – Ministério da Economia e do Emprego

NUT – Nomenclatura de Unidades Territoriais

PME – Pequena(s) e Média(s) Empresa(s)

ÍNDICE DE TABELAS E FIGURAS

Tabela 1. Variáveis específicas à empresa que influenciam a sobrevivência das empresas.	6
Tabela 2. Modelo empírico - amostra geral.....	19
Tabela 3. Tabela de Vida.....	32
Tabela 4. Tabela de Vida com os estimadores Kaplan-Meier e de Nelson-Aalen.....	34
Tabela 5. Teste Log-rank.....	37
Tabela 6. Estimções semi-paramétricas e paramétricas aplicadas aos fundadores.	39
Tabela 7. Estimções semi-paramétricas e paramétricas aplicadas aos gestores.	41
Tabela 8. Estimções semi-paramétricas e paramétricas aplicadas aos trabalhadores.	42
Figura 1. Tipos de Censura.....	25
Figura 2. Função de sobrevivência.	33
Figura 3. Curva de sobrevivência de Kaplan-Meier.....	35
Figura 4. Curvas de sobrevivência de Kaplan-Meier entre grupos.....	36

CAPÍTULO I - INTRODUÇÃO

1.1. Contextualização e motivação

O estudo da sobrevivência das empresas tem motivado os economistas ao longo de vários anos e para o caso português foram já realizados vários estudos relacionados com esta temática. Geroski, Mata e Portugal (2010), Baptista, Karaöz e Mendonça (2007), Mata e Portugal (2002 e 1994), Mata (1996) e Mata, Portugal e Guimarães (1995) foram alguns dos autores que se debruçaram sobre este tipo de estudos tendo os resultados gerais sido semelhantes. Como principais conclusões podemos apontar que se o fundador esteve anteriormente no setor no qual cria a sua empresa há um impacto positivo sobre a capacidade de sobrevivência da mesma (Baptista, Karaöz e Mendonça, 2007). Outros fatores como a dimensão da empresa e o crescimento da empresa e do setor podem ser apontados como influenciadores da sobrevivência (Geroski, Mata e Portugal, 2007, Mata, 1996, Mata, Portugal e Guimarães, 1995 e Mata e Portugal, 1994).

Porém, estes estudos carecem de uma análise mais pormenorizada da influência que os colaboradores têm sobre o desempenho da organização sendo que o impacto do capital humano apenas é analisado nos dois estudos mais recentes. A presente dissertação procura assim colmatar esta falha na literatura analisando a influência do capital humano na sobrevivência das empresas. Para o efeito são construídos três índices de capital humano, um para os patrões, outro para os gestores e um último para os restantes trabalhadores, que visam caracterizar a força de trabalho em termos de experiência e educação. Além do referido índice, são ainda incluídas na análise outras características específicas a cada indivíduo, nomeadamente a experiência no setor no qual a empresa é fundada e a experiência dos gestores em cargos similares.

É neste contexto que este trabalho se propõe centrar contribuindo para a literatura que estuda em geral a sobrevivência das empresas e, em particular, para a literatura mais recente e crescente que se tem debruçado sobre o contributo dos colaboradores na dinâmica empresarial.

Assim, espera-se que as conclusões alcançadas permitam que sejam feitas recomendações fundamentadas a empresários e decisores políticos no âmbito da definição de políticas económicas.

1.2. **Objetivos e questões de investigação**

Segundo Wooldridge (2010), o objetivo da maioria dos estudos empíricos em ciências sociais consiste em analisar a forma como uma ou várias variáveis afetam outra. No caso do presente trabalho, o objetivo geral baseia-se na análise da sobrevivência das empresas portuguesas e dos fatores que a influenciam e, em particular, responder à seguinte questão de investigação: *Qual o impacto do capital humano na sobrevivência das empresas?*.

Para cumprir tal objetivo pretende-se realizar um estudo empírico – cujo evento de interesse é a duração da empresa, i.e., o tempo que empresa sobrevive no mercado desde que nasce até à sua morte. Por outras palavras, pretendemos estimar as funções de sobrevivência das empresas observadas.

Para tal, foram utilizados dados da base de dados ‘Quadros de Pessoal’ do Ministério do Trabalho e da Solidariedade Social, recolhidos para o período compreendido entre 1999 e 2007, tendo sido analisadas 1120 empresas.

1.3. **Estrutura da dissertação**

O presente documento está estruturado da seguinte forma. O capítulo 2 apresenta a revisão de literatura que resume as conclusões de outros autores no que concerne ao impacto do capital humano na sobrevivência das empresas bem como a identificação de alguns determinantes da mesma. No capítulo 3 é apresentada a metodologia, fontes utilizadas para obtenção de dados, os critérios utilizados para a seleção das empresas e é descrito o modelo empírico. O capítulo 4 apresenta o modelo econométrico e os resultados empíricos. Por último, no capítulo 5 são apresentadas as conclusões deste estudo, apontadas algumas recomendações assim como as limitações e sugestões para investigação futura.

CAPÍTULO II - ENQUADRAMENTO TEÓRICO

2.1. Introdução

Este capítulo tem por objetivo apresentar uma revisão da literatura sobre sobrevivência das empresas e capital humano. Para o efeito organizou-se o capítulo da seguinte forma. Na secção 2.2 é analisada a relação entre o capital humano e sobrevivência das empresas. Na secção 2.3 são analisados outros determinantes da sobrevivência específicos à empresa. Na secção 2.4 são apontados os determinantes externos, nomeadamente setoriais bem como alguns fatores macroeconómicos importantes na explicação da sobrevivência das empresas. Por último, na secção 2.5 são retiradas as conclusões.

2.2. Capital humano e sobrevivência das empresas

“Joseph Schumpeter e Theodore Schultz demonstraram, em meados do século 20, que os recursos humanos são um fator de produção muito mais importante que os recursos naturais” (Piazza-Georgi, 2002: 462). Além disso, entidades como clientes e investidores necessitam, segundo Bosma et al. (2004), de avaliar as empresas com base em características observáveis que se assumem ter uma relação com a qualidade da empresa e o capital humano do fundador é apontado como uma dessas características.

O capital humano – que pode ser definido como “stock de capacidades pessoais que os agentes económicos têm à sua disposição” (Piazza-Georgi, 2002: 463) e não sendo apenas “o resultado da educação formal, mas incluindo também a experiência e a aprendizagem prática que se aplica no trabalho” (Davidsson e Honig, 2003: 306) é um ativo complexo e difícil de analisar. Segundo Helfat e Peteraf (2003), cada indivíduo possui, pelo menos, capital humano e social. O primeiro com total incidência sobre as capacidades, conhecimentos, educação e experiência e o segundo relacionado com os laços sociais com outros elementos. Ambos, de acordo com Baptista et al. (2007), desempenham um papel importante no sucesso das novas empresas. No que concerne ao tipo de capital devemos ainda fazer a distinção proposta por Brüderl et al. (1992) entre capital geral e capital específico. O capital geral é medido pela escolaridade enquanto o capital específico inclui a experiência numa indústria particular.

Segundo vários autores, entre os quais Gimmon e Levie (2010), Colombo e Grilli (2005) e Bates (1990), as características do empreendedor estão diretamente relacionadas com a

longevidade e desempenho da empresa com os quais o capital humano do empreendedor tem uma relação positiva (Bates, 1990). Por outras palavras, o histórico de carreira e a experiência dos fundadores no mercado de trabalho são determinantes da estratégia corporativa e do sucesso (Preisendörfer e Voss, 1990: 110). Esta ideia é também defendida por Bosma et al. (2004: 6) segundo o qual quanto maior o capital do fundador, maior deverá ser a produtividade da empresa.

Outra característica do empreendedor também relevante é a sua educação. De acordo com Bates (1990), a educação do empreendedor é um dos principais determinantes do montante de empréstimos bancários concedidos para a formação do negócio. Esta perspetiva também é corroborada por Saridakis, Mole e Storey (2008) que reconhecem a importância destas duas variáveis afirmando que estas promovem a sobrevivência das empresas. Mais recentemente, e contrapondo a posição defendida anteriormente, Gimmon e Levie (2010) defenderam que o efeito da educação sobre a capacidade de sobrevivência da empresa não é significativo.

Além da educação do fundador, a sua experiência determina fortemente o sucesso das novas organizações (Dahl e Reichstein, 2007) e leva a uma maior sobrevivência (Colombo e Grilli, 2005). Com a criação de uma *spin-off*¹, o fundador pode transportar grande parte do conhecimento adquirido – o que os consumidores procuram, quem são os fornecedores, quem são os seus concorrentes, qual a melhor tecnologia, etc. - para a sua empresa. Então, no que concerne ao capital humano do empreendedor pode-se acrescentar que, tal como é dito por Bosma et al. (2004), o facto de o indivíduo possuir experiência na indústria na qual cria a sua empresa tem um efeito positivo sobre todas as medidas de desempenho e a sua experiência em liderança tem o mesmo efeito sobre o tempo de sobrevivência.

Segundo Siegel et al. (1993), além da experiência anterior na indústria por parte do fundador, este tipo de experiência de toda a equipa de gestão é relevante até porque é através desta variável que as empresas de baixo e de elevado crescimento se distinguem. Ainda no que concerne à experiência do fundador, deve-se salientar o contributo de Gimmon e Levie (2010) valorizando-a relativamente à gestão de negócios anteriores. O seu passado faz com que sejam obtidas melhores capacidades técnicas e de gestão, que seja criada uma rede de contactos e que seja mais fácil o acesso a determinada informação (Shane, 2000). Contudo, é necessário ter em atenção se o fundador esteve sempre presente em empresas sobreviventes dado que “a

¹ Empresas que são fundadas por grupos ou pessoas originárias da mesma área laboral, ou de uma empresa existente (Eriksson e Kuhn, 2006).

experiência do fundador numa empresa que saiu do mercado é uma desvantagem direta para a nova empresa” (Dahl e Reichstein, 2007: 13).

Ainda no que concerne ao fundador, a idade deve ser uma variável a considerar dado que, segundo Preisendörfer e Voss (1990), a sua relação com o tempo de sobrevivência da empresa é positiva. Contudo, segundo Thompson (2007), se ao medirmos o impacto da idade do fundador controlarmos para a experiência pré-entrada verificamos que a idade não influencia a sobrevivência da empresa.

É possível que, com o crescimento da empresa, o empreendedor não consiga coordenar todas as atividades diárias sendo obrigado a delegar funções. Sob esta perspetiva importa analisar o impacto das características dos gestores, por exemplo, testar se as empresas geridas por indivíduos sem experiência têm a mesma capacidade de sobrevivência que aquelas cuja gestão está a cargo de indivíduos com experiência anterior na função. Algumas empresas empregam gestores sem qualificação e outras, não reconhecendo a importância que uma boa gestão tem nos seus resultados, não contratam nenhum profissional para essa área podendo este ser um determinante do encerramento das unidades empresariais. O tecido empresarial português é maioritariamente constituído por micro e PMEs², onde os empreendedores, além de proprietários, são também os seus gestores, o que intensifica a importância do seu capital humano.

A análise do capital humano está, como foi já referido, aliada a um considerável grau de complexidade não existindo uma variável direta que o consiga medir. Assim sendo, apesar de existirem diferentes tipos de capital humano, na generalidade dos estudos empíricos este é medido através dos anos de escolaridade, excluindo outras componentes importantes. Segundo Portela (2001), podemos construir um índice de capital humano que inclui diferentes dimensões da produtividade do indivíduo, nomeadamente, a escolaridade, experiência e competências não observáveis diretamente.

2.3. Determinantes da sobrevivência específicos à empresa

Pela análise da literatura conclui-se que existe uma vasta lista de variáveis relacionadas com as características das empresas que influenciam a sua probabilidade de sobrevivência ainda que todas concorram em mercados com recursos escassos. Isto acontece porque as empresas

² “Em 2009, existiam em Portugal 348 552 PME, representando 99.7% do total de sociedades do setor não financeiro e cerca de 59% do volume de negócio.” (INE, 2011)

defrontam um processo de seleção natural em que umas evoluem com o tempo adaptando-se às mudanças exigidas no ambiente em que se inserem. Outras empresas não o conseguem fazer e acabam por extinguir-se. Contudo, muitos foram os autores que concluíram que existe uma vasta lista de possíveis fatores, relacionados não só com o ambiente mas também com o empreendedor e com a empresa, que podem impedir que as empresas vingam.

Assim sendo, para além das variáveis centrais a este estudo, importa destacar outras consideradas relevantes na literatura com vista a controlar para este tipo de efeitos. A Tabela 1 apresenta uma síntese da literatura no que respeita às variáveis de controlo relacionadas com as características das empresas apontadas com mais frequência como influenciadoras da sua sobrevivência e mortalidade.

Tabela 1. Variáveis específicas à empresa que influenciam a sobrevivência das empresas.

<i>Variável</i>	<i>Autor</i>
Dimensão	Kaniovski e Peneder (2008)
	Manjón-Antolin e Arauzo-Carod (2008)
	Cefis e Marsili (2006)
	Cefis e Marsili (2005)
	Honjo (2000)
	Audretsch (1995)
	Audretsch e Mahmood (1995)
	Mata, Portugal e Guimarães (1995)
	Mata e Portugal (1994)
	Jovanovic (1982)
Localização	Strotmann (2007)
	Stearns, Reynolds e Williams (1995)
Crescimento	Kanivski e Peneder (2008)
	Mata e Portugal (2002)
	Ilmakunnas e Topi (1999)
	Audretsch (1995)
	Phillips e Kirchhoff (1989)

Fonte: Do autor.

É vasta a lista de autores que defendem que as empresas podem incrementar a sua capacidade de sobrevivência através da dimensão podendo esta relação decorrer do facto de as empresas de menor dimensão terem capacidade financeira mais limitada. Esta medida pode ser representada pelo número de estabelecimento que quanto mais elevado for, maior será a probabilidade de sobrevivência (Manjón-Antolin e Arauzo-Carod, 2008). Tal pode dever-se ao

facto de para uma empresa com vários estabelecimentos sair do mercado ter que assistir a uma falha de todos os seus estabelecimentos ou, pelo menos, que o sucesso de alguns deles não seja suficiente para compensar a falha de outros.

Segundo Strotmann (2007), ao contrário do que se podia esperar, as empresas localizadas em áreas urbanas têm menor probabilidade de sobrevivência do que as empresas localizadas em áreas rurais. A justificação para esta conclusão apresentada por Stearns, Reynolds e Williams (1995) prende-se com o facto de as áreas urbanas serem marcadas por uma concorrência elevada cuja abundância de recursos não consegue colmatar. Por outro lado, as áreas rurais oferecem a possibilidade de serem explorados nichos em que a concorrência é baixa. No que concerne ao crescimento, as empresas com maior probabilidade de sobrevivência apresentam maiores taxas de crescimento até pela redução dos seus custos médios (Mata e Portugal, 2002). Contudo, apesar desta relação, não se pode dizer que o crescimento leva à sobrevivência nem que as empresas que sobrevivem irão crescer.

2.4. Determinantes setoriais

Espera-se que a concentração tenha um efeito negativo sobre a sobrevivência das novas empresas tal como é defendido por Holmes, Stone e Braidford (2010) e Wagner (1994). Esta relação pode ser explicada pelo facto de “maiores taxas de concentração levarem a uma maior probabilidade de as incumbentes se envolverem numa maior retaliação contra as novas empresas” (Wagner, 1994: 145). É, então, importante testar se uma maior concorrência leva inevitavelmente a uma maior luta por parte das empresas – especialmente das novas empresas – para se manterem no mercado (Audretsch e Mahmood, 1994). A resposta a este teste pode não ser tão simples como parece pois o conluio³ é mais fácil quando existem mais empresas e, com esta união poderá ser mais fácil as empresas envolvidas sobreviverem (Geroski, Mata e Portugal, 2010).

Além dos determinantes relacionados com as características das empresas e dos determinantes setoriais, o impacto de algumas variáveis ligadas ao ambiente no qual a empresa se insere deve ser tido em conta. Entre essas variáveis macroeconómicas interessa destacar o crescimento de mercado que tende a estar positivamente relacionado com a sobrevivência das empresas (Kaniowski e Peneder, 2008, Manjón-Antolin e Arauzo-Carod, 2008, Audretsch e

³ Diferentes tipos de acordos estabelecidos entre empresas que visam reduzir a concorrência.

Mahmood, 1994 e Wagner, 1994). Isto porque, segundo Wagner (1994: 146), “parece mais fácil pequenas empresas crescerem em indústrias prósperas do que em indústrias nas quais, pela perda de quota de mercado, as concorrentes podem retaliar”. Esta justificação vai ao encontro da justificação dada por Mata e Portugal (1994: 233) de que “em indústrias de rápido crescimento pode ser mais fácil sobreviver, desde que as empresas possam crescer sem infligir perdas de quota de mercado às suas rivais e, portanto, a probabilidade de reações agressivas é mais baixa”. Porém, este impacto não reúne consenso entre os autores analisados sendo que Honjo (2000) defende que as taxas de saída tendem a estar positivamente relacionadas com o crescimento de mercado e Audretsch e Mahmood (1995) não encontram evidência de que o crescimento de mercado afeta a sobrevivência das empresas.

Em suma, no decorrer dos trabalhos analisados, neste trabalho busca-se testar se:

Hipótese 1: Existe uma relação positiva entre a educação e a experiência dos fundadores a sobrevivência das empresas.

Hipótese 2: Existe uma relação positiva entre a idade dos fundadores e a sobrevivência das empresas.

Hipótese 3: Existe uma relação positiva entre a educação e a experiência dos gestores e a sobrevivência das empresas.

Hipótese 4: Existe uma relação positiva entre a educação e a experiência dos restantes trabalhadores e a sobrevivência das empresas.

2.5. **Conclusão**

Neste capítulo foi apresentada uma revisão de literatura sintetizando os principais fatores que determinam a sobrevivência empresarial, dando relevância ao capital humano. Genericamente, é sugerido que quanto maior for este capital maior será a sobrevivência das empresas.

Além do capital humano foram apresentados outros determinantes relacionados com as características das próprias empresas, do setor e do ambiente no qual se inserem, ou seja, foram também incluídos fatores de contexto. Percebe-se que há uma relativa concordância quanto aos vários fatores que devem ser incluídos na análise da sobrevivência das empresas. No entanto, podemos verificar que, por um lado, existe maior consenso na literatura no que respeita aos efeitos das variáveis como a experiência do fundador, a dimensão, a localização e o nº de estabelecimentos detidos pela empresa e a concorrência enfrentada. Por sua vez, as conclusões auferidas pelos autores variam para o impacto da idade e da educação do fundador e do crescimento de mercado.

Realçando, mais uma vez, a complexidade ligada à medição do capital humano importa referir que a maioria dos autores não demonstra, de forma clara, como este é medido. É relevante a inclusão de todas as vertentes de capital humano referidas neste capítulo devendo, para tal, as variáveis serem pensadas de forma a responder a esta exigência.

De seguida, no capítulo 3, é apresentada a fonte de dados, a descrição das variáveis utilizadas e o modelo empírico.

CAPÍTULO III - DADOS, VARIÁVEIS E MODELO EMPÍRICO

3.1. Introdução

O presente capítulo tem por objetivo apresentar a metodologia de investigação pensada para responder à questão central deste estudo, concretamente: - *Qual o impacto do capital humano na sobrevivência das empresas?* Para o efeito debruça-se sobre a descrição das fontes dos dados, as variáveis empíricas e o modelo empírico, deixando para o capítulo seguinte a apresentação do modelo econométrico e os resultados das estimações.

Na secção 3.2 é apresentada a fonte dos dados utilizados, discutidas as vantagens e desvantagens e a definição da amostra. Na secção 3.3 é apresentado o modelo empírico e é descrita a operacionalização das variáveis utilizadas.

3.2. Fontes e definição das amostras

Para alcançar os objetivos propostos, além da investigação documental, foi realizado um estudo empírico cujos dados constam numa base amplamente utilizada no âmbito da economia industrial e da economia do trabalho. Esta corresponde aos ‘Quadros de Pessoal’ do Ministério do Trabalho e da Solidariedade Social que decorre de um questionário anual obrigatório a todas as empresas que operam em Portugal desde 1982⁴.

Os Quadros de Pessoal constituem uma base de dados em painel na medida em que várias empresas, respetivos trabalhadores e estabelecimentos, são acompanhados em vários anos. Durante o período abrangido – 1985 a 2009 – apenas não está disponível informação completa para os anos 1991 e 2001. Esta ausência de dados bem como o facto de nem todas as empresas estarem presentes em todos os anos faz com que este seja um painel não balanceado⁵.

Esta base de dados é preponderante para estudos relacionados com o mercado de trabalho português visto que contém informação sobre uma vasta lista de empresas, seus estabelecimentos e colaboradores⁶. No grupo de dados correspondentes à empresa encontram-se variáveis como a localização, o volume de vendas, o ano de fundação e o número de

⁴ À data deste estudo estavam disponíveis, na sala de Quadros de Pessoal da Escola de Economia e Gestão da Universidade do Minho, dados de 1985 a 2009.

⁵ Se a base contemplasse dados para todos os anos de 1985 a 2009 estávamos perante um painel balanceado.

⁶ A administração pública, os trabalhadores independentes e o serviço doméstico não são cobertos pelo inquérito.

empregados. Quanto aos trabalhadores é possível, entre outras variáveis, analisar a sua situação profissional, remuneração e idade e, por último, no que concerne aos estabelecimentos, estão presentes, por exemplo, a localização e o setor. Uma outra característica que interessa apontar prende-se com o facto de, inicialmente, o mês de Março ser usado como mês de referência tendo, a partir de 1994, dado lugar ao mês de Outubro. Apesar desta alteração, é perfeitamente possível cruzar dados anteriores e posteriores a 1994 sem que a análise seja severamente afetada. Este cruzamento é possível dada a dimensão longitudinal desta base e a identificação que é atribuída a empresas, trabalhadores e estabelecimentos. Podem ser apontadas, de acordo com Mata e Portugal (1994), algumas outras vantagens decorrentes desta base de dados sendo algumas delas a sua abrangência e a possibilidade de cruzamento entre dados da empresa, dados dos estabelecimentos e dados dos trabalhadores. Tal como todos os dados em painel, esta base é mais favorável que os dados seccionais - relativos a um único momento do tempo – na medida em que permite, por exemplo, obter uma “estimação mais eficiente dos parâmetros (...), reduz os problemas de identificação” (Verbeek, 2000: 311), controla a heterogeneidade⁷ e diminui a colinearidade⁸.

No entanto, algumas limitações podem ser enumeradas, nomeadamente, o facto de “não podermos distinguir as fusões e aquisições de verdadeiras mortes”, Mata e Portugal (2002: 12). Isto acontece porque perante uma operação deste tipo, o identificador de uma das empresas permanece e o da(s) restante(s) desaparece da base de dados. Além disso, na contabilização dos trabalhadores são excluídos os trabalhadores a recibos verdes que na atualidade têm uma forte presença no mercado laboral português.

Dada a elevada fiabilidade que os dados apresentam para o ano de 1999 este foi escolhido como o ano base, ou seja, todas as empresas selecionadas foram fundadas neste ano⁹. A escolha deste ano está também ligada ao facto de este ser, segundo Conraria, Alexandre e Pinho (2010: 3), o início de uma mudança de regime económico com a adesão de Portugal ao euro e a fixação do câmbio que constitui “provavelmente o mais importante acontecimento na economia portuguesa das duas últimas décadas.”

Além da definição do ano de fundação de 1999 como o ano em análise, várias decisões foram tomadas ao longo da construção da amostra. Uma delas passou pela resolução do

⁷ Resulta da omissão de variáveis importantes e a sua não inclusão pode levar a estimadores inconsistentes.

⁸ Este termo é aplicado quando há uma forte correlação entre duas ou mais variáveis independentes.

⁹ Esta identificação é feita através da variável que nos dá o ano de constituição da empresa, *ancon*.

problema relacionado com a existência de empresas intermitentes (ausentes da base de dados por um ano mais que uma vez) tendo-se optado pela eliminação destas empresas.

Feita esta seleção e criadas as variáveis a utilizar neste estudo, verificou-se que das 7079 empresas que permaneceram na amostra após a aplicação dos critérios enunciados, apenas 2379 têm fundadores identificados, 3422 têm gestores e 6420 têm trabalhadores que não são gestores nem fundadores. O último procedimento para a seleção da amostra passou pela manutenção apenas das empresas que têm informação para todas as variáveis criadas. Desta forma, restaram 1120 empresas que formam o *risk set*, isto é, o conjunto de empresas que estão em risco de ocorrência do evento em cada ponto do tempo (Allison, 1984: 16).

3.3. Modelo empírico e variáveis empíricas

Uma vez selecionadas as empresas fundadas em 1999 foram reunidos os dados das mesmas até ao ano 2007¹⁰ para que o acompanhamento das mesmas fosse possível. Procedeu-se ao cruzamento dos dados destas empresas com os dos seus trabalhadores o que permitiu chegar às observações necessárias para criação das variáveis.

Interessa agora especificar o modelo empírico a estimar. Para tal, foi seguida a revisão de literatura na identificação dos principais determinantes da sobrevivência das empresas e definido um modelo de sobrevivência onde esta é função de variáveis específicas às empresas, aos trabalhadores e ao ambiente no qual estas se inserem. Assumindo que a sobrevivência da empresa i é dada por Y_i o modelo empírico pode ser representado da seguinte forma:

$$Y_i = \alpha + X_i\beta + Z_i\lambda + A_i\delta + \varepsilon_i \quad (3.1)$$

em que X é o vetor correspondente às características das empresa, Z é o vetor das características específicas aos trabalhadores, A corresponde ao vetor das características do ambiente no qual a empresa se insere e ε representa o termo de erro do modelo. α , β , λ e δ são os vetores que representam os parâmetros a estimar.

As variáveis utilizadas para estimar cada um dos efeitos não foram, na sua totalidade, retiradas diretamente da base de dados, são sim o resultado de variáveis construídas, trabalhadas e recodificadas a partir das originais para melhor se adaptarem a este estudo (ver

¹⁰ Apesar de estarem disponíveis, à data da dissertação, dados até ao ano 2009 o último ano utilizado foi 2007. Isto decorre do facto de um dos critérios de identificação da morte das empresas ser a sua ausência por dois anos consecutivos não sendo este possível de aplicar nos anos 2008 e 2009. Além disso, não foi possível acompanhar a empresa em 2001 porque não temos dados para esse ano. Uma alternativa seria a imputação de dados, contudo, a imputação tem limitações e, por norma, são usados apenas os dados originais (StataCorp, 2011).

quadro A.1 em anexo). De seguida, é então descrita a construção das variáveis sendo estas divididas em quatro categorias: variáveis ligadas ao capital humano, às empresas, ao ambiente no qual estas se inserem - sendo as variáveis das duas últimas categorias usadas para efeitos de controlo – e variáveis necessárias à análise de duração.

Capital Humano

Com vista à construção do Índice de Capital Humano dos trabalhadores – *ICH_trabalhadores* - sugerido por Portela (2001), cuja expressão é dada por

$$\begin{aligned}
 ICH_i = meduc * & \left(0.5 + \frac{e^{\frac{educ_i - meduc}{educ}}}{1 + e^{\frac{educ_i - meduc}{educ}}} \right) \\
 & * \left(0.5 + \frac{e^{\frac{exper_i - mexper}{sexper}}}{1 + e^{\frac{exper_i - mexper}{sexper}}} \right) \\
 & * \left(0.5 + \frac{e^{\frac{efeito_i - mefeito}{sefeito}}}{1 + e^{\frac{efeito_i - mefeito}{sefeito}}} \right)
 \end{aligned} \tag{3.2}$$

, foram criadas três variáveis que nos dão a educação dos trabalhadores, a experiência, o efeito específico e respetivas médias e desvios-padrão. Na base de dados não está disponível informação sobre a educação formal dos trabalhadores pelo que foi necessário criar uma aproximação da mesma através da variável que nos dá as habilitações escolares - *habil* (3 dígitos) – (ver quadro A.2 em anexo). No que diz respeito à experiência, e porque também não há na base de dados dos Quadros de Pessoal uma variável direta com a experiência dos trabalhadores, foi usada a experiência potencial = idade – educação – 6. Assume-se que a escolaridade começa aos 6 anos de idade, a entrada do trabalhador no mercado de trabalho acontece logo após serem concluídos os estudos, os indivíduos não trabalham e estudam ao mesmo tempo e nunca estão em situações de desemprego (trabalham sem interrupções desde o final dos estudos até à reforma). Por último, o efeito específico do trabalhador foi gerado pelo método de efeitos fixos¹¹ através da seguinte regressão salarial:

¹¹ Portela (2011) recomenda o método de efeitos fixos sendo esta recomendação comprovada pelo teste de Hausman (ver anexo A.3).

Com este modelo é possível controlar os efeitos das variáveis omitidas que variam entre indivíduos mas permanecem constantes ao longo do tempo. Assim sendo, neste modelo as variáveis que não variam no tempo não podem ser consideradas como variáveis explicativas sendo absorvidas pela interseção da reta de regressão.

$$\begin{aligned} \lnsalario = & \beta_0 + \beta_1 x educ + \beta_2 x educ^2 + \beta_3 x exper + \beta_4 x exper^2 + \beta_5 x antig + \beta_6 x \\ & antig^2 + \beta_7 x educexper + \beta_8 x ano + \beta_9 x IHH + \beta_{10} x nest + \varepsilon \end{aligned} \quad (3.3)$$

em que *lnsalario* corresponde ao logaritmo do salário que, por sua vez, é dado pela soma entre as remunerações bases e as remunerações regulares, *educ2* e *exper2* respeitam à quadrática da educação e da experiência respetivamente, *antig* corresponde à antiguidade e *antig2* à sua quadrática, *educexper* dá a interação entre a educação e a experiência, *ano* representa uma dummy para cada ano, *IHH* é o índice de Herfindahl- Hirschman¹² no ano em questão e, por último, *nest* corresponde ao número de estabelecimentos.

Com o intuito de incluir o impacto que os gestores podem ter sobre a capacidade de sobrevivência das empresas foi necessário identificar os indivíduos que ocupam cargos de gestão (ver quadro A.4 em anexo). Feito isto, pode ser criada uma variável que contabiliza o número de gestores existentes em cada empresa e analisar o seu impacto, contudo, mais importante que o número de gestores acredita-se que sejam as qualificações dos mesmos. Então, um ICH semelhante ao descrito anteriormente foi construído para os gestores – *ICH_gestor*. Assim sendo foram geradas, de forma semelhante ao que foi feito para os trabalhadores, as variáveis necessárias ao cálculo do ICH: educação, experiência e efeito específico dos gestores, bem como as respetivas médias e desvios-padrão.

Para os fundadores o método de criação do ICH difere do utilizado para os gestores e para os restantes trabalhadores. Dado que os salários não são fornecidos para grande parte dos fundadores optou-se por incluir dois efeitos específicos. Para tal foi criado um painel no qual os fundadores são acompanhados em anos anteriores e posteriores e no qual constam os salários, se nesse ano o fundador não era fundador, e o volume de vendas da empresa na qual se encontrava, se nesse ano o fundador já o era. Feito isto, através dos seguintes modelos de efeitos fixos

$$\lnsalario = \beta_0 + \beta_1 x exper_fundador + \beta_2 x i.indat1 + \beta_3 x i.ano + \varepsilon \quad (3.4)$$

$$\lnvendas = \beta_0 + \beta_1 x exper_fundador + \beta_2 x i.indat1 + \beta_3 x i.ano + \varepsilon \quad (3.5)$$

¹² Supondo que existem M empresas o IHH é dado por $IHH = \sum_{i=1}^M \left[\frac{Q_i}{\sum_{i=1}^M Q_i} \right]^2$ em que Q é a quota de mercado. Quando maior o IHH maior é a concentração de mercado, ou seja, se $IHH=1$ estamos perante uma situação de monopólio.

em que *exper_fundador* corresponde à experiência geral dos fundadores, *inda1* à identificação o setor de atividade (ver quadro A.5 em anexo), e *ano* ao ano em questão, foi possível obter dois efeitos específicos, *efeito_vendas* e *efeito_salarios*.

O ICH dos fundadores é calculado de formas distintas:

$$\begin{aligned}
 ICH_i = meduc * & \left(0.5 + \frac{e^{\frac{educ_i meduc}{educ}}}{1 + e^{\frac{educ_i educ}{seduc}}} \right) \\
 * & \left(0.5 + \frac{e^{\frac{exper_i - mexper}{sexper}}}{1 + e^{\frac{exper_i - mexper}{sexper}}} \right) \\
 * & \left(0.5 + \frac{e^{\frac{efeito_salario_i - mefeito_salario}{sefeito_salario}}}{1 + e^{\frac{efeito_salario_i - mefeito_salario}{sefeito_salario}}} \right) \\
 * & \left(0.5 + \frac{e^{\frac{efeito_vendas_i - mefeito_vendas}{sefeito_vendas}}}{1 + e^{\frac{efeito_vendas_i - mefeito_vendas}{sefeito_vendas}}} \right)
 \end{aligned} \tag{3.6}$$

para os fundadores para os quais dispomos do *efeito_salario* e do *efeito_vendas*. Caso só esteja disponível o efeito através das vendas, apenas esse é incluído sendo que o mesmo acontece para o efeito obtido através dos salários.

Para os fundadores para os quais não estão disponíveis nem o *efeito_vendas* nem o *efeito_salario* o ICH é calculado da seguinte forma:

$$\begin{aligned}
 S_i = meduc * & \left(0.5 + \frac{e^{\frac{educ_i - meduc}{seduc}}}{1 + e^{\frac{seduc_i - seduc}{seduc}}} \right) \\
 * & \left(0.5 + \frac{e^{\frac{exper_i - mexper}{sexper}}}{1 + e^{\frac{exper_i - mexper}{sexper}}} \right)
 \end{aligned} \tag{3.7}$$

Sendo o capital humano o ponto principal da presente dissertação mereceu cuidado acrescentado e, além das variáveis anteriores, várias foram as variáveis criadas. Entre elas estão as correspondentes à experiência dos fundadores, dos gestores e dos restantes trabalhadores no

setor no qual a empresa foi fundada – *expset_fundador*, *expset_gestor* e *expset_trabalhador*, respetivamente. Para construir estas variáveis foi feito o acompanhamento dos indivíduos nos anos anteriores¹³ à sua entrada na empresa sendo que a experiência resulta da soma do número de anos em que estes estiveram no mesmo setor.

A variável que mede a experiência do fundador em cargos de gestão – *expgest_fundador* –, bem como a correspondente à experiência em empresas que não sobreviveram – *exp_mort* – foram criadas da mesma forma. Contudo, estas variáveis acabaram por não ser incluídas no estudo dada a sua reduzida variabilidade que levava a inferências erradas. No caso dos gestores foi também analisada a sua experiência em cargos similares – *expcargo_gestor*.

Ainda no que concerne às variáveis relacionadas com o fundador foi analisado o impacto da sua idade – *idade_fundador*.

Empresa

A dimensão da empresa – *dimensao* – pode ser medida de várias formas sendo uma delas o número de estabelecimentos que é retirado diretamente da base de dados original. Relativamente às variáveis relacionadas com as empresas foram também criadas variáveis de crescimento, nomeadamente o logaritmo do crescimento do número de trabalhadores e do crescimento do volume de vendas – *cr_trabalhadores* e *cr_vendas*. No caso do volume de vendas esta é uma variável que existe na base de dados e que não necessita de qualquer transformação mas, por outro lado e apesar de também existir na base, o número de trabalhadores resultou da sua contabilização e não da variável existente no ficheiro das empresas que nos fornece esse mesmo número. Esta escolha resulta do facto de a variável original – *pemp* – conter incoerências para algumas empresas.

Ambiente

Relativamente à região na qual a empresa se insere foi usada a variável representativa da região NUTs II – *localizacao*. Para testar o impacto da concentração industrial sobre a sobrevivência das empresas foi criada a variável *IHH* que corresponde ao logaritmo do Índice de Herfindahl-Hirschman aquando da fundação. Este índice resulta da soma dos quadrados das quotas de mercado das empresas que operam no mercado em questão quotas essas que podem ser calculadas através das vendas visto que representam as vendas de determinada

¹³ Os dados correspondentes a 1985 não foram usados porque apresentam várias falhas.

empresa em percentagem das vendas totais do mercado em que atua. Pode assumir valores entre $\frac{1}{n}$ e 1 sendo que quanto mais elevado for o valor maior é o grau de concentração. Interessa também testar o impacto do crescimento do setor que pode ser medido através do número de empresas. Para tal foi criada a variável *cr_empresas* sendo esta o logaritmo do crescimento absoluto.

Duração

O evento de interesse (ver explicação do conceito na secção 4.3) corresponde à morte das empresas tendo então sido gerada a variável *died* que corresponde a uma variável dummy igual a 1 se a empresa morre e igual a 0 caso contrário. Para a sua construção foram aplicadas as sugestões do manual Eurostat e OCDE (2007). Assim sendo, uma empresa é considerada morta no ano N se não surge na base de dados no ano N+1 e no ano N+2. Esta metodologia foi já anteriormente avançada por Mata, Portugal e Guimarães (1995: 463), autores de referência nesta temática, que defendem que “as empresas cuja saída ocorre num único momento são consideradas vivas nesse período (...). Pelo contrário, as empresas cuja saída temporária é maior que 1 são excluídas da base de dados.” De acordo com Mata e Portugal (1994: 229), “uma empresa pode não ser incluída nos ficheiros num determinado ano por diversas razões além do encerramento permanente da sua operação. Isto pode dever-se ao facto da empresa suspender a sua operação, porque nesse ano não pagou aos trabalhadores ou simplesmente porque não preencheu o inquérito. Podemos reduzir o risco de contar estas empresas entre as mortes classificando como saída apenas aquelas que não surgem nos ficheiros por 2 anos consecutivos.” Se uma empresa não tiver nenhum trabalhador remunerado deve também ser considerada morta.

A par destes critérios constantes no referido manual, e apesar da recomendação dada por Mata e Portugal (1994), optou-se por eliminar os casos em que as empresas, mesmo estando ausentes da base de dados apenas por um ano, tal se verifica mais que uma vez sendo estas consideradas empresas intermitentes. Seguindo estes passos podiam acontecer ‘reentradas’, ou seja, uma empresa podia ser considerada morta num ano e, caso no ano seguinte cumprisse todos os requisitos enunciados, voltava a estar presente na base de dados. Optou-se então por excluir informações posteriores à primeira morte de cada empresa.

A duração da empresa - *tempo* – é dada pela diferença entre o ano presente e o ano de constituição sendo esta a variável dependente. A esta diferença foi adicionado um ano para que

a empresa que nasce e morre no mesmo ano tenha duração igual a 1. Por último, a variável censura - *status* – é uma variável dummy que assume valor 1 se a empresa morre e valor 0 se a empresa continua viva no final do período.

Na Tabela 2 são apresentadas a variável dependente e as variáveis independentes e respetivas estatísticas descritivas.

Tabela 2. Modelo empírico - amostra geral¹⁴.

<i>Variável</i>	<i>Descrição</i>	<i>Média</i>	<i>D.P.</i>	<i>Min.</i>	<i>Máx.</i>
tempo	Duração da empresa	6.562	3.123	1.000	9.000
ICH_trabalhador	ICH dos trabalhadores	2.143	0.434	1.065	4.250
ICH_fundador	ICH dos fundadores	3.052	0.948	1.066	7.797
ICH_gestor	ICH dos gestores	2.909	0.896	1.285	7.537
expset_trabalhador	Experiência dos trabalhadores no setor	2.070	2.002	0.000	11.000
expset_gestor	Experiência dos gestores no setor	2.345	2.783	0.000	15.500
expset_fundador	Experiência dos fundadores no setor	0.010	0.203	0.000	6.125
expcargo_gestor	Experiência dos gestores em cargos similares	0.739	1.151	0.000	7.333
idade_fundador	Idade do fundador	37.835	9.525	18.000	70.000
dimensao	Número de estabelecimentos	1.043	0.265	1.000	7.000
IHH	Logaritmo do IHH aquando da fundação	0.006	0.009	0.002	0.238
cr_vendas	Logaritmo do crescimento absoluto das vendas da empresa	9.917	1.314	6.153	14.812
cr_empresas	Logaritmo do crescimento absoluto do número de empresas por setor	23.967	1.006	19.776	24.942
localizacao	Região NUTs II	14.883	3.587	11.000	30.000

Fonte: Cálculos próprios baseados nos Quadros de Pessoal, MEE.

3.4. Conclusão

No capítulo que aqui se encerra foi apresentada a fonte de dados utilizada para responder à questão “*Qual o impacto do capital humano sobre a sobrevivência das empresas?*”. Esta base consiste nos Quadros e Pessoal do MEE e contempla informação sobre empresas, trabalhadores e estabelecimentos da economia portuguesa. A riqueza desta base de dados passa, além da vasta informação contemplada, pela possibilidade de cruzamento entre empresas e trabalhadores que em muito contribui para este estudo.

Para cumprir os objetivos propostos foi retirada informação dos ficheiros das empresas e dos trabalhadores indo ao encontro do que é referido na literatura. Numa segunda parte foi apresentado o modelo empírico com a descrição das variáveis utilizadas e a operacionalização

¹⁴ Inclui todas as empresas independentemente da categoria de PAVITT à qual pertencem

das mesmas. Foram também apresentadas algumas estatísticas descritivas para as variáveis utilizadas.

No capítulo seguinte é apresentado o modelo econométrico e respetivos resultados.

CAPÍTULO IV - RESULTADOS

4.1. Introdução

Neste capítulo é apresentado o modelo econométrico implementado com base nas variáveis empíricas apresentadas no capítulo anterior. São também condensados os resultados da investigação a fim de concluir quanto ao impacto dos vários fatores que se acreditam poder influenciar a sobrevivência das empresas e feito o cruzamento com o contributo dos vários autores descrito na secção da revisão de literatura

O capítulo é encerrado com uma pequena síntese dos impactos encontrados.

4.2. Modelo Econométrico

Os dados recolhidos serão tratados com recurso à análise quantitativa, em particular serão utilizadas técnicas econométricas, cuja escolha foi determinada pelas características das amostras finais dos dados.

O modelo utilizado deverá ser capaz de estimar a probabilidade das empresas sobreviverem e este será formulado com base no trabalho de vários autores, entre os quais Carreira e Teixeira (2011), Dahl e Reichstein (2007), Fox (2002), Honjo (2000), Shane (2000), Mata, Portugal e Guimarães (1995), Audretsch e Mahmood (1994) e Wagner (1994).

Para a análise de duração – também designada por análise de sobrevivência ou análise de transição – vários modelos podem ser utilizados. Tal como é apontado por Audretsch e Mahmood (1995), vários autores, entre os quais Evans (1987), Hall (1987) e Phillips e Kirchoff (1989), mesmo tendo acesso a dados longitudinais não estimaram o modelo de duração visto que os dados estavam agrupados em intervalos temporais de 5 anos. Porém, no caso da base de dados dos Quadros de Pessoal são disponibilizados dados de caráter anual.

Do modelo de duração constam alguns conceitos fundamentais que importa conhecer. Exemplo disso são a função densidade de probabilidade – $f(t)$ –, a função de sobrevivência – $S(t)$ –, a função de risco – $h(t)$ – e a função de risco cumulativa – $H(t)$.

A função densidade de probabilidade pode ser apresentada da seguinte forma

$$f_{(t)} = \lim_{\Delta t \rightarrow 0} \frac{\Pr(t \leq T < t + \Delta t)}{\Delta t} \quad (4.1)$$

ou seja, corresponde ao limite da probabilidade de um indivíduo morrer no intervalo compreendido entre t e Δt .

A função de sobrevivência corresponde à probabilidade de uma empresa sobreviver além de determinado período t dado que sobreviveu até esse momento. Então, em termos probabilísticos é dada por

$$S_{(t)} = P(T > t) = 1 - \Pr[T \leq t] = 1 - F_{(t)} = \int_t^{\infty} f(x)dx \quad (4.2)$$

em que T é tempo entre o evento inicial e o evento de interesse (fundação e morte da empresa), $P(T>t)$ corresponde à probabilidade da empresa não morrer até um determinado momento t dado que sobreviveu até aí e $F_{(t)}$ é a função distribuição cumulativa.

Uma vez que a função sobrevivência corresponde a uma probabilidade varia entre 0 e 1 assumindo valor 1 no momento inicial ($t=0$) e é decrescente até 0 à medida que t aumenta (Cleves et al., 2010).

$$\frac{\partial S}{\partial t} < 0 \quad (4.3)$$

Por exemplo, $S_{(3)} = \Pr(T>3)$ diz respeito à probabilidade de uma empresa sobreviver pelo menos quatro anos dado que já sobreviveu três anos.

Trabalhando a função de sobrevivência chega-se a

$$F_{(t)} = 1 - S_{(t)} = \Pr(T \leq t) \quad (4.4)$$

que define a função de distribuição cumulativa dos tempos de duração que pode ser interpretada como a probabilidade de uma empresa não sobreviver após o momento t /probabilidade de duração até ao tempo t .

A função de risco – também designada por taxa condicional de falha -, apresentada por Kiefer (1988), dá-nos a probabilidade condicionada¹⁵ calculada como sendo a probabilidade de determinado evento ocorrer no instante t dado que tal não aconteceu até esse momento. Esta função é dada pela expressão seguinte

¹⁵ A probabilidade condicional corresponde, por exemplo, à probabilidade de uma empresa morrer aos 4 anos dado que viveu até essa idade enquanto a probabilidade não condicionada mede simplesmente a probabilidade das empresas morrerem até aos 4 anos. Então, é a probabilidade de um evento – morte – acontecer supondo uma condição anterior.

$$h_{j(t)} = h_{0(t)} \exp (X_{(t)}\beta) \quad (4.5)$$

na qual h_j é taxa de risco enfrentada pela empresa j , h_0 diz respeito à função de risco base que depende da idade das novas empresas, X representa as covariáveis/variáveis explicativas¹⁶ e β é vetor dos parâmetros da regressão.

Vários autores concordam com a forma exponencial da função de risco, porém, dado que não há concordância quanto à forma da função de risco base, optou-se por não modelar essa mesma função.

Esta função pode também ser definida como o rácio entre o limite da probabilidade de uma empresa falhar no intervalo compreendido entre t e Δt dado que sobreviveu até esse momento e a variação do tempo.

$$h_{j(t)} = \lim_{\Delta t \rightarrow 0} \frac{\Pr (t + \Delta t > T \geq t | T \geq t)}{\Delta t} \quad (4.6)$$

O rácio entre a função densidade de probabilidade e a função de sobrevivência é uma outra forma de representar esta função. Então, segundo Lancaster (1990):

$$h_{j(t)} = \frac{f_{(t)}}{S_{(t)}} = \frac{f_{(t)}}{1 - F_{(t)}} = 1 - \bar{F}_{(t)} \quad (4.7)$$

Por último, a função de risco pode ser o resultado da razão entre duas funções de sobrevivência cada uma respeitante a um momento dividida pela extensão do intervalo:

$$h_{(t)} = \frac{S_{(t_1)} - S_{(t_2)}}{(t_2 - t_1)S_{(t_2)}} \quad (4.8)$$

Em suma, segundo Cox e Oakes (1984), a função de risco – também conhecida como taxa de falha instantânea, taxa de mortalidade condicional ou simplesmente força de mortalidade – representa a forma como a probabilidade instantânea de falha varia com o tempo. Se o tempo

¹⁶ “Determinantes internos e externos que se assumem influenciar a taxa de risco” (Carreira e Teixeira, 2008). Dada a inclusão de variáveis explicativa esta função é utilizada em modelos semi-paramétricos ou paramétricos (ver secções 4.3.2. e 4.3.3.).

for medido em anos e o risco calculado for de 1,6 isto significa que num ano a morte deverá ocorrer 1,6 vezes.

De acordo com Cleves et al. (2010) e Royston e Lambert (2011), podemos ainda apresentar a função de risco cumulativa que mede o risco total que foi acumulado até ao momento t :

$$H_{(t)} = \int_0^t h(u) du \quad (4.9)$$

através da qual podemos chegar a

$$H_{(t)} = \int_0^t \frac{f(u)}{S(u)} du = - \int_0^t \frac{1}{S(u)} \left\{ \frac{d}{du} S(u) \right\} du = - \ln\{S(t)\} \quad (4.10)$$

No decorrer desta expressão podemos desenvolver restantes da seguinte forma

$$S_{(t)} = \exp\{-H_{(t)}\} \quad (4.11)$$

que nos dá a relação entre a função de sobrevivência e a função de risco cumulativa.

$$F_{(t)} = 1 - \exp\{-H_{(t)}\} \quad (4.12)$$

$$f_{(t)} = h_{(t)} \exp\{-H_{(t)}\} \quad (4.13)$$

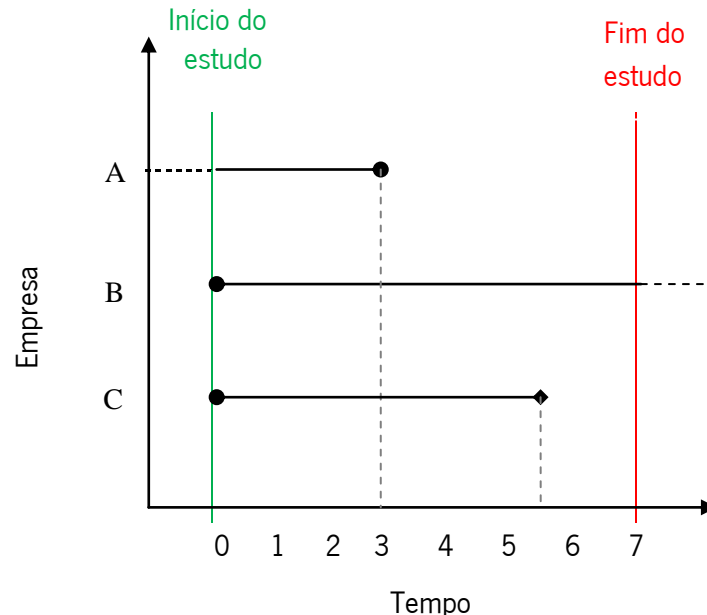
Uma vez definidas as funções base em análise de duração interessa definir a data de origem/evento inicial, a escala de tempo e o evento final/evento de interesse/evento falha. O evento inicial consiste em “alguma mudança qualitativa que ocorre num ponto específico do tempo” (Allison, 1984: 9) e, neste estudo, corresponde ao ano de fundação da empresa, o tempo é medido em anos e o evento de interesse – que corresponde ao momento em que a empresa deixa de estar no estado em que se encontrava – é a morte da mesma.

Independentemente da tipologia do modelo, este deve permitir que seja avaliado o comportamento duma variável resposta não negativa, ou seja, deve possibilitar o estudo das empresas e do tempo que decorre entre o evento inicial e o evento final conhecido como o

tempo de falha. Contudo, tal como é referido por Holmes, Stone e Braidford (2010), o principal problema que surge quando analisamos dados de duração prende-se com os dados censurados que são, de certa forma, dados perdidos. Isto ocorre para as empresas que não morrem durante o período de estudo - censura à direita - e, para estes casos, podemos apenas falar de um tempo de sobrevivência censurado. Ao estabelecer os períodos de análise, todas as empresas que tinham morrido antes do início do período não são observados - *censura à esquerda* - sendo esta censura também aplicada às empresas para as quais não conhecemos o seu momento inicial. Porém, dado que para o presente estudo foram seleccionadas empresas com base no seu ano de fundação não estão presentes dados censurados à esquerda. Uma terceira forma de censura prende-se com o facto de ser difícil saber exactamente em que momento o evento de falha ocorre - censura por intervalo, ou seja, podemos verificar que a empresa no momento t está viva e no momento $t+1$ já não está o que nos permite saber que o evento de falha ocorreu em algum momento entre t e $t+1$ não sendo possível analisar exactamente qual foi esse momento.

Na Figura 1 são ilustrados estes três tipos de censura.

Figura 1. Tipos de Censura.



Fonte: Do autor.

A empresa A ilustra uma situação de censura à esquerda, a censura à direita é representada pela empresa B e, por sua vez, a empresa C corresponde a um caso de censura por intervalo.

A eliminação das empresas que respeitam estes critérios poderia levar a resultados enviesados pelo que as ditas observações censuradas devem ser incluídas na análise estatística da seguinte forma através da variável de censura:

$$\delta_i = \begin{cases} 1 & \text{se } t_i \text{ é o tempo de falha} \\ 0 & \text{se } t_i \text{ é o tempo de censura} \end{cases}$$

Devemos utilizar um “modelo estatístico capaz de acomodar estas durações incompletas” (Mata e Portugal, 2002: 13) e, para tal, devem ser usados procedimentos paramétricos e não-paramétricos (Mata e Portugal, 1994). Nas secções seguintes são apresentados os vários modelos que podem ser aplicados em estudos deste tipo entre os quais estão os modelos não-paramétricos – que não incluem qualquer variável explicativa -, os modelos semi-paramétricos – que contemplam uma parte paramétrica (variáveis explicativas) e uma parte não paramétrica (constante) - e os modelos paramétricos –que relacionam a variável dependente com vários parâmetros.

4.2.1. **Análise não paramétrica**

Segundo Cleves et al. (2010), perante a inexistência de covariáveis ou quando estas são qualitativas podem ser aplicados métodos não paramétricos para estimar a função sobrevivência. De seguida, são resumidos três destes métodos, nomeadamente, as tabelas de vida, o estimador de Kaplan-Meier e o estimador de Nelson-Aalen.

Tabelas de Vida

As tabelas de vida são, segundo Audretsch e Mahmood (1994), um método não paramétrico que, tal como defende Wagner (1994), pode ser utilizado para encontrar as taxas de risco. Com este método o horizonte temporal é dividido em intervalos facilitando desta forma os cálculos necessários. Contudo, dado o reduzido número de intervalos, estes cálculos não são muitos utilizados pelo que é recomendado o uso do estimador apresentado de seguida.

Kaplan-Meier

Um estimador não paramétrico da função sobrevivência ou da função de risco que é usualmente aplicado é o estimador de Kaplan-Meier ou Produto Limite que consiste num método de máxima verosimilhança. Uma vantagem acrescida deste estimador prende-se com o fato de apenas necessitar da variável tempo dispensando a inclusão de variáveis explicativas. Este

estimador, resume a contabilização do número de observações, do número de saídas e da probabilidade condicional de sobrevivência em cada um dos tempos.

Perante a existência de dados censurados e tempos de sobrevivência que variam entre 1 e m este é dado por

$$\hat{S}_{(t)} = \prod_{j:t_j \leq t} \left(\frac{n_j - \delta_j}{n_j} \right) = \prod_{j:t_j \leq t} \left(1 - \frac{\delta_j}{n_j} \right), \text{ para } t_1 \leq t \leq t_m \quad (4.14)$$

de acordo com o qual em t_j existiam n_j empresas em risco mas que não morreram e δ_j que morreram. Então, para o tempo t , é feita a soma de todos os tempos de sobrevivência inferiores a t – sendo que para cada um deles é feito o cálculo da probabilidade condicionada de sobreviver até t_{j+1} dado que sobreviveu até t_j – e feitas as respetivas multiplicações, ou seja, é calculado o produto para todos os períodos de risco desde a fundação da empresa até ao momento t .

Nelson-Aalen

Um outro estimador não paramétrico utilizado neste tipo de análise é o estimador de Nelson-Aalen dado por

$$\hat{H}_{(t)} = \sum_{j|t_j \leq t} \frac{dj}{n_j} \quad (4.15)$$

sendo este bastante utilizado quando o interesse recai sobre as taxas de risco cumulativas.

No decorrer dos objetivos do presente estudo, estes estimadores não são suficientes na medida em que a análise do efeito das covariáveis sobre a sobrevivência das empresas é importante dado que empresas com a mesma data de fundação podem enfrentar diferentes funções de risco no decorrer das suas diferentes características. Para que estas sejam consideradas devemos recorrer a modelos paramétricos ou semi-paramétricos.

4.2.2. Análise semi-paramétrica

Com recurso à análise semi-paramétrica os estimadores anteriores são complementados indo ao encontro daquilo que se pretende.

Dentro dos modelos semi-paramétricos o modelo de riscos proporcionais¹⁷ de Cox é o mais popular tendo sido utilizado por vários autores entre os quais Audretsch e Mahmood (1994 e 1995), Audretsch (1995), Mata et al. (1995), Agarwal e Audretsch (2001), Disney et al. (2003) e Cefis e Marsili (2005). Este modelo mostra-se mais vantajoso em relação aos modelos não paramétricos na medida em que permite a inclusão de variáveis explicativas – regressores esses que podem justificar as diferentes durações das empresas (Lancaster, 1990) - e isola o efeito de cada variável controlando a heterogeneidade. A designação semi-paramétrica está relacionada com o facto de apenas as covariáveis serem tratadas parametricamente.

Este modelo de duração é dado pelo logaritmo da probabilidade da empresa encerrar em t visto que sobreviveu até $t-1$, explicado pelas variáveis X . A função de risco é dada por

$$h_j(t) = h_0(t) \exp(\beta'X) = h_0(t) \exp(\beta_0 + X_j\beta_x) \quad (4.16)$$

onde $h_j(t)$ é a função de risco de morte da empresa j , $h_0(t)$ corresponde à função de risco base¹⁸ e X são as variáveis explicativas que correspondem à parte paramétrica do modelo. Sendo assim, esta expressão dá-nos, segundo Cleves et al. (2010), o risco enfrentado pela empresa j que depende do risco enfrentado por todas as empresas modificado por X . A interpretação dos parâmetros β não é mensurável sendo apenas possível verificar se as variáveis têm um impacto positivo ou negativo sobre a variável resposta. A presença da exponencial faz com que os riscos estimados nunca sejam negativos ($0 \leq h(t, X) < \infty$).

Dividindo a expressão por $h_0(t)$ chegamos à função das razões de risco (Cox e Oakes, 1984):

$$\frac{h(t/X_1, X_2, \dots, X_k)}{h_0(t)} = \exp(\beta_1 X_1 + \beta_2 X_2 + \dots + \beta_k X_k) \quad (4.17)$$

Uma limitação do modelo de riscos proporcionais de Cox prende-se com o facto de assumir que os riscos são proporcionais sendo que a função de risco é a mesma para todas as empresas e as variações das variáveis explicativas levam a deslocamentos paralelos da função de risco base. Então, antes da aplicação do modelo interessa implementar um teste da proporcionalidade do risco.

¹⁷ Imaginando que as empresas com determinadas características, como por exemplo o facto de pertencerem a um setor específico, são monitoradas durante 12 anos e se verifica que o risco delas morrerem é o dobro do risco das empresas que pertencem a outros setores, a suposição da proporcionalidade dos riscos requer que o rácio de risco seja o mesmo no primeiro ano, no segundo, no décimo segundo ou em qualquer outro momento do período de análise.

¹⁸ Dado que $h(t;X)=h_0(t)$ quando $X=0$. Nesta função os valores das variáveis explicativas correspondem aos valores médios sendo, então, a componente de risco comum a todos os indivíduos cuja definição é arbitrária. Por outras palavras, é o risco de uma empresa que cumpre as condições padrão (Cox e Oakes, 1984).

Teste de Resíduos de Schoenfeld

O teste proposto por Schoenfeld (1982) é dado pela diferença entre o valor da covariável da empresa que morre menos o seu valor esperado. Este teste define os resíduos como sendo independentes do tempo, então, o i -ésimo resíduo pode ser comparado com t_i para testar a suposição de riscos proporcionais.

Para este teste podemos formular, segundo Cleves et al. (2010), as seguintes hipóteses:

H_0 : O logaritmo da função de risco é constante ao longo do tempo.

Perante a possibilidade de aplicação do modelo de Cox podemos definir que a variável tida como dependente é o tempo de sobrevivência (longevidade) da empresa e relativamente às covariáveis a revisão de literatura efetuada sugere quatro tipos que importa incluir na análise. Concretamente, variáveis específicas ao capital humano da equipa de gestão, dos fundadores e dos restantes trabalhadores, variáveis correspondentes à empresa, variáveis ligadas ao setor e, por último, ao ambiente em que insere a empresa. Uma técnica bastante utilizada neste tipo de estudo para decidir qual o modelo que deverá ser utilizado e quais as variáveis que devem ser incluídas é a técnica dos resíduos de Cox e Snell (1968). Esta técnica consiste na construção das curvas de relação de resíduos estimados em cada um dos modelos e da comparação dos mesmos. O modelo escolhido deverá ser aquele que apresenta uma curva de resíduos mais próxima da curva resultante do modelo não-paramétrico.

4.2.3. Análise paramétrica

Na categoria dos modelos paramétricos incluem-se as distribuições Exponencial, Weibull, Log-Normal, Log-Logística e Gama generalizada entre outras.

Exponencial

As expressões que interessam para o presente estudo são dadas, na distribuição exponencial, da seguinte forma

$$f(t) = \alpha \exp(-\alpha t) \quad (4.18)$$

$$S(t) = \exp(-\alpha t) \quad (4.19)$$

$$F(t) = 1 - \exp(-\alpha t) \quad (4.20)$$

$$h(t) = h_{0(t)} \exp(X_{(t)}\beta) \quad (4.21)$$

Segundo Royston e Lambert (2011), este modelo tem funções risco que podem crescer, decrescer ou ser constantes.

Weibull

A distribuição Weibull pode ser vista como uma generalização da distribuição exponencial e as funções chave em análises de duração são dadas por

$$f(t) = \alpha \gamma t^{\gamma-1} e^{-(\alpha t)^\gamma} \quad (4.22)$$

$$S(t) = e^{-(\alpha t)^\gamma} \quad (4.23)$$

$$F(t) = 1 - e^{-(\alpha t)^\gamma} \quad (4.24)$$

$$h(t) = \alpha \gamma (\alpha t)^{\gamma-1} \quad (4.25)$$

Segundo Royston e Lambert (2011), no modelo Weibull a função risco vai sempre na mesma direção com o tempo (para cima, para baixo, ou constante). Apesar de muitos dados terem riscos que atingem um pico e depois decrescem, por exemplo, esta distribuição tem uma ampla aplicabilidade e adequa-se a uma elevada variabilidade de formas.

Log-normal

A distribuição log-normal é muito utilizada dado que o logaritmo do tempo segue uma distribuição normal com média μ e desvio-padrão σ facilitando a interpretação.

Neste caso as funções são dadas por

$$f(t) = \frac{1}{t\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\log t - \mu}{\sigma} \right)^2 \right] \quad (4.26)$$

$$S(t) = 1 - \theta \left(\frac{\log t - \mu}{\sigma} \right) \quad (4.27)$$

$$F(t) = \theta \left(\frac{\log t - \mu}{\sigma} \right) \quad (4.28)$$

$$h(t) = \frac{\frac{1}{t\sigma\sqrt{2\pi}} \exp \left[-\frac{1}{2} \left(\frac{\log t - \mu}{\sigma} \right)^2 \right]}{1 - \theta \left(\frac{\log t - \mu}{\sigma} \right)} \quad (4.29)$$

em que θ é a função distribuição cumulativa da distribuição normal.

A função de risco não é monótona, é crescente até alcançar um pico sendo depois decrescente.

Uma desvantagem do modelo log-normal é a sensibilidade perante pequenos tempos de falha.

Gamma generalizado

A distribuição gama generalizada inclui as distribuições Weibull, Exponencial e Log-normal.

(4.30)

$$f(t) = \frac{\gamma \alpha^{k\gamma}}{\Gamma(k)} t^{k\gamma-1} e^{-(\alpha t)^\gamma}$$

$$S(t) \begin{cases} I(k(\alpha t)^\gamma, \alpha) & \text{se } \gamma < 0 \\ 1 - I(k(\alpha t)^\gamma, \alpha) & \text{se } \gamma > 0 \end{cases} \quad (4.31)$$

em que

$$I(k, x) = \frac{1}{\Gamma(k)} \int_0^x x^{k-1} dx \quad (4.32)$$

$$F(t) = 1 - S(t) \quad (4.33)$$

$$h(t) = \frac{\frac{1}{\alpha \Gamma(k)} \left(\frac{t}{\alpha}\right)^{k-1} e^{-\frac{t}{\alpha}}}{1 - I\left(k, \frac{t}{\alpha}\right)} \quad (4.34)$$

Log-logística

A representação das funções na distribuição log-logística é a seguinte

$$f(t) = \frac{\alpha \gamma t^{\gamma-1}}{(1 + \alpha t^\gamma)^2} \quad (4.35)$$

$$S(t) = \frac{1}{1 + \alpha t^\gamma} \quad (4.36)$$

$$F(t) = 1 - \frac{1}{1 + \alpha t^\gamma} \quad (4.37)$$

$$h(t) = \frac{\gamma \alpha t^{\gamma-1}}{1 + \alpha t^\gamma} \quad (4.38)$$

Gompertz

Esta distribuição é aplicável a dados com taxas de risco que aumentam ou diminuem exponencialmente com o tempo (Cleves et al., 2010 e Royston e Lambert, 2011).

$$f(t) = \alpha \exp[\gamma t - \alpha \gamma^{-1}(e^{\gamma t} - 1)] \quad (4.39)$$

$$S(t) = \exp[-\gamma^{-1} \exp(\beta_0 + X_j \beta_x) \{ \exp(\gamma t) - 1 \}] \quad (4.40)$$

$$F(t) = 1 - \exp[-\gamma^{-1} \exp(\beta_0 + X_j \beta_x) \{ \exp(\gamma t) - 1 \}] \quad (4.41)$$

$$h(t) = \exp(\gamma t) \exp(\beta_0 + X_j \beta_x) \quad (4.42)$$

Para escolher qual o modelo mais adequado deve ser aplicada a técnica de Cox-Snell (1968).

4.3. Resultados

Nesta secção serão apresentados os resultados dos vários tipos de modelo aplicados começando pelos não paramétricos, seguindo-se os resultados dos modelos semi-paramétricos e terminando com os modelos paramétricos e respetivos testes.

4.3.1. Modelos não paramétricos

Numa fase inicial foram aplicados modelos não paramétricos, nomeadamente as tabelas de vida, o estimador de Kaplan-Meier e o estimador de Nelson-Aalen. Na Tabela 3 são apresentados os resultados de um dos instrumentos mais utilizados para quantificar o perfil de sobrevivência da população em estudo: tabela de vida.

Tabela 3. Tabela de Vida.

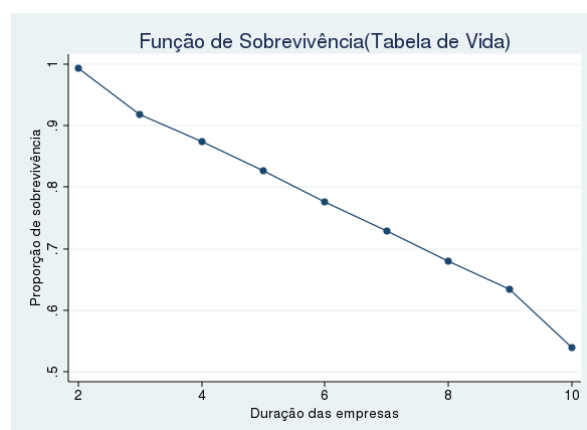
Intervalo de tempo		Obs. início do intervalo	Nº de mortes	Nº de empresas perdidas	Sobrevivência	Erro padrão	[95% Int. Conf.]	
1	2	1120	8	0	0.993	0.003	0.989	0.996
2	3	1112	83	0	0.919	0.008	0.901	0.933
3	4	1029	50	0	0.874	0.010	0.853	0.892
4	5	979	54	0	0.826	0.011	0.802	0.847
5	6	925	55	0	0.777	0.012	0.751	0.800
6	7	870	54	0	0.729	0.013	0.702	0.754
7	8	816	55	0	0.680	0.014	0.651	0.706
8	9	761	50	8	0.635	0.014	0.606	0.662
9	10	703	57	646	0.540	0.017	0.506	0.572

Fonte: Cálculos próprios baseados nos Quadros de Pessoal, MEE.

Nas primeiras duas colunas da tabela é apresentada a idade intervalar¹⁹, as observações presentes na terceira coluna dizem respeito ao número de empresas em risco, ou seja, o número de empresas vivas no início de cada período e, por outro lado, o número de empresas que morrem em cada período é apresentado na quarta coluna. Na coluna 5 é visível o número de empresas censuradas à direita verificando-se que no final do período existia um número razoável de empresas – 646 das 1120 – que não tinham ainda desaparecido. As 8 empresas censuradas em 2006 (intervalo de 8 a 9) são aquelas que estiveram ausentes da base de dados em 2007 mas voltam a estar presentes em 2008 e 2009 não sendo esta ausência indicativa de morte. Relembrando os critérios utilizados para definição da morte das empresas, um deles estipula que uma empresa não é considerada morta quando está ausente da base de dados por um ano sendo condição obrigatória que isso só aconteça uma vez. Na construção da tabela de vida, e porque 2007 é o último ano de acompanhamento, estes casos são encarados como empresas perdidas no tempo na construção da tabela de vida.

Por último, dado que na coluna seis são apresentadas as taxas de sobrevivência e uma vez que estas correspondem a taxas do final do período verifica-se, por exemplo, que cerca de 99% das empresas fundadas em 1999 sobreviveram dois anos ou mais. Entre o oitavo e o nono ano de atividade, estima-se que aproximadamente 37% das empresas encerram. Estima-se também que apenas 45% das empresas permaneçam ativas por mais de 10 anos. A probabilidade de sobrevivência sofre uma queda gradual à medida que a duração aumenta. A Figura 2 corresponde à representação gráfica da tabela de vida.

Figura 2. Função de sobrevivência.



Fonte: Cálculos próprios baseados nos Quadros de Pessoal, MEE.

¹⁹ O intervalo é aberto à direita, ou seja, no primeiro intervalo são incluídas as empresas que morrem com um ano de idade e não as que morrem com dois, essas são incluídas no segundo intervalo.

A figura anterior permite uma observação mais clara relativamente à evolução da sobrevivência das empresas ao longo do período analisado. É visível um acentuado decréscimo na proporção de empresas sobreviventes verificando-se que esta tende para zero à medida que o tempo avança.

Também na estimação não paramétrica foram obtidas funções de sobrevivência pelo método de Kaplan-Meier e pelo método de Nelson-Aalen sendo estas apresentadas na Tabela 4.

Tabela 4. Tabela de Vida com os estimadores Kaplan-Meier e de Nelson-Aalen.

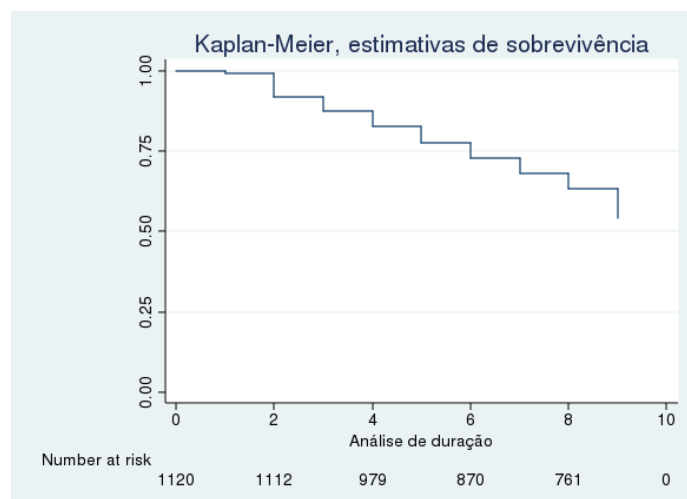
<i>Tempo</i>	<i>Função de Sobrevivência de Kaplan-Meier</i>	<i>Função de risco cumulativa de Nelson-Aalen</i>
1	0.993	0.007
2	0.919	0.082
3	0.874	0.130
4	0.826	0.186
5	0.777	0.245
6	0.729	0.307
7	0.680	0.375
8	0.634	0.442
9	0.543	0.585

Fonte: Cálculos próprios baseados nos Quadros de Pessoal, MEE.

Na segunda coluna da tabela está o estimador de Kaplan-Meier que nos dá a probabilidade de sobrevivência em determinado momento usando a proporção de empresas sobreviventes. Este estimador inclui as observações censuradas assumindo que essas empresas têm uma probabilidade de sobrevivência igual àquelas para as quais se conhece a data do evento de interesse. Na terceira coluna apresenta-se o estimador de Nelson –Aalen que nos dá a função de risco cumulativa. Dado que esta mede o risco de ocorrência da morte da empresa até ao instante t, pelo resultado ilustrado na segunda linha da Tabela 4, por exemplo, pode-se concluir que o risco total acumulado até ao momento 2 é de cerca de 8%. Deve ser destacado o facto de mais de 20% das empresas morrerem no quinto período e de menos de 60% das empresas sobreviverem após 9 anos.

Segundo Hosmer e Lemeshow (1999: 57), em estudos de sobrevivência, devemos apresentar gráficos com o estimador de Kaplan-Meier. Assim sendo, na Figura 3 é ilustrada a probabilidade de uma empresa permanecer em atividade em cada momento do tempo.

Figura 3. Curva de sobrevivência de Kaplan-Meier.



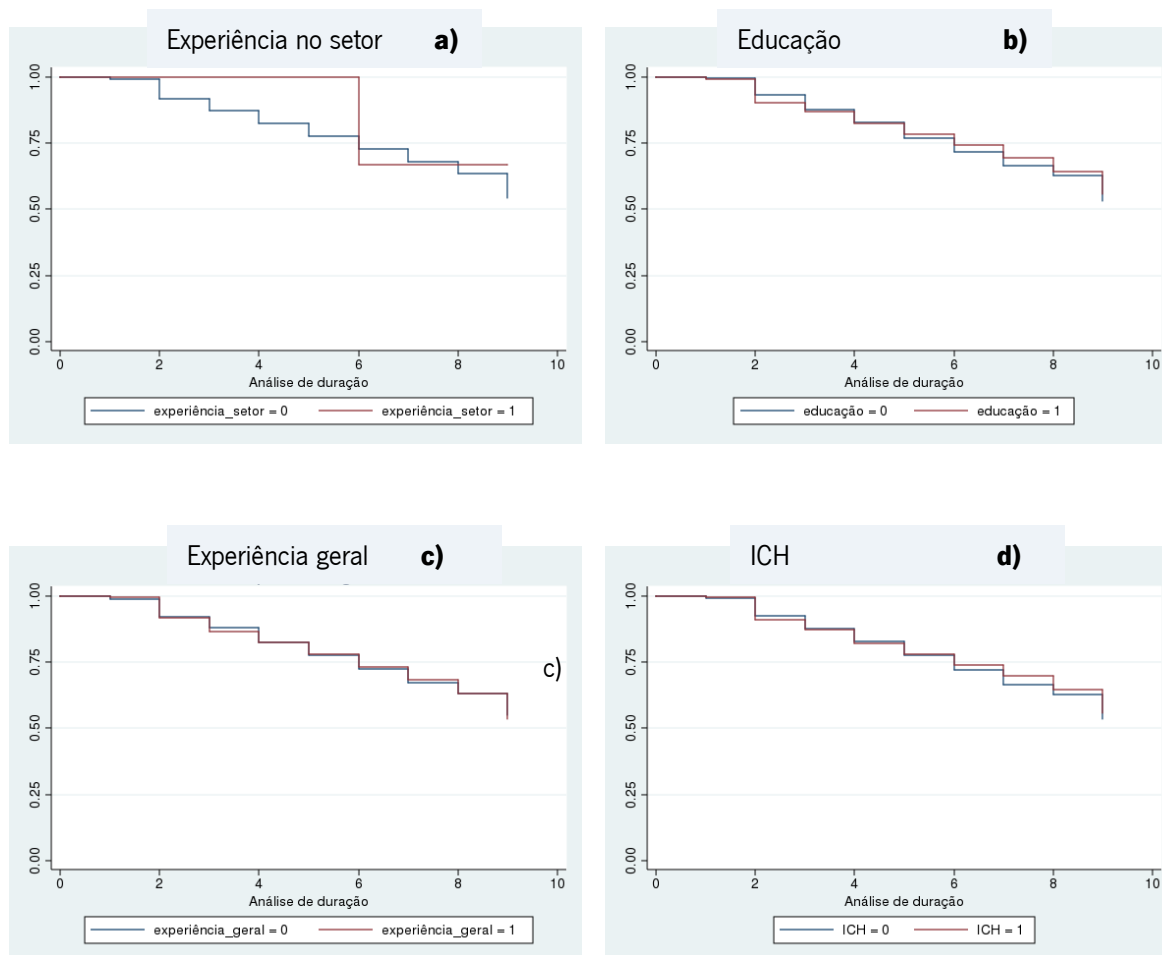
Fonte: Cálculos próprios baseados nos Quadros de Pessoal, MEE.

Tendo sido aplicado o estudo às empresas fundadas em 1999, e sendo o ano de fundação o indicador do momento inicial, no momento $t=0$ todas as empresas estavam vivas e, por isso, a função de sobrevivência é 1 no início do período. A maior queda no número de empresas ativas acontece no primeiro ano sendo registadas quedas relativamente mais suaves daí em diante. Ao fim de 6 anos existiam apenas 870 empresas em atividade tendo este número diminuído para 761 ao fim de 8 anos. Tal como se tinha verificado em estimações anteriores, a função de sobrevivência é, então, negativamente inclinada, tendendo para 0 ao longo do tempo. Contudo, neste caso e dado que existem observações censuradas no final do período, a função nunca será igual a 0. Tal só aconteceria no momento em que todas as empresas deixassem de existir, contudo, não é utilizado um intervalo de tempo suficientemente alargado para poder comprovar esse facto.

Este estimador pode ser também ser aplicado para comparar empresas de diferentes grupos através de variáveis qualitativas. Para esse efeito, foram geradas as médias de algumas variáveis relacionadas com os fundadores e variáveis dummy que distinguem se a empresa está acima ou abaixo da média. De forma mais clara e dando como exemplo o caso da educação do fundador, foi gerada a sua média medida entre todas as empresas independentemente do setor, da região ou de outras características. Feito isto, foi gerada uma variável dummy – *educação* – que assume valor 1 se a educação dos fundadores da empresa é superior à média e que assume valor 0 se for inferior. O mesmo foi feito para a experiência geral do fundador – *experiência_geral* –, experiência do fundador no setor no qual a empresa é fundada –

experiência_setor - e para o ICH dos fundadores - *ICH*. As estimações feitas com estas variáveis pelo método de Kaplan-Meier são apresentadas na Figura 4.

Figura 4. Curvas de sobrevivência de Kaplan-Meier entre grupos.



Fonte: Cálculos próprios baseados nos Quadros de Pessoal, MEE.

Dado que as linhas a vermelho são representativas das empresas acima da média e as linhas a azul representam as empresas abaixo da média, para todas as variáveis é visível uma igualdade na probabilidade de sobrevivência no início do período. Pelo gráfico a), que consta da Figura 4 pode verificar-se que a probabilidade de sobrevivência diminui ao longo do tempo sendo, entre o sexto e o oitavo ano mais baixa, ao contrário do que se esperava, para as empresas cujos fundadores têm experiência no setor no qual as empresas são fundadas. Por

sua vez, as empresas cujos fundadores têm um ICH superior à média têm probabilidade de sobrevivência ligeiramente superior, a partir do sexto ano, como é visível pelo gráfico d).

A experiência e a educação são as medidas de capital humano mais utilizadas nos estudos analisados sendo as suas médias a base de construção dos gráficos b) e c). O gráfico b) ilustra uma proximidade acentuada entre as curvas da sobrevivência das empresas cujos fundadores têm educação superior e inferior à média. Contudo, a probabilidade de sobrevivência é ligeiramente superior para as empresas cujos fundadores têm educação superior à média sendo que isto só acontece de forma clara a partir do sexto ano. No gráfico c) é visível que as empresas cujos fundadores têm uma maior experiência geral têm maior probabilidade de sobrevivência.

Porém, para confirmar a significância destas diferenças torna-se necessário implementar um teste estatístico adequado à comparação entre curvas de sobrevivência de diferentes grupos. Um possível teste é o teste *Log-rank* cujas hipóteses são:

$$\begin{cases} H_0: \text{não há diferença entre as curvas} \\ H_1: \text{há diferença entre as curvas} \end{cases}$$

Os resultados do teste são apresentados na Tabela 5.

Tabela 5. Teste Log-rank.

<i>Variável</i>	<i>Prob>chi2</i>
Educação dos fundadores	0.438
Experiência geral	0.697
Experiência dos fundadores no setor no qual a empresa é criada	0.682
ICH dos fundadores	0.480

Fonte: Cálculos próprios baseados nos Quadros de Pessoal, MEE.

Dadas as hipóteses testadas, para um nível de significância de 5%, quando *p-valor* < 0.05 existe diferença significativa entre as curvas de sobrevivência dos grupos em questão. Para nenhuma das variáveis a hipótese nula é rejeitada o que significa que não há diferença estatisticamente significativa entre as curvas de sobrevivência das empresas que estão acima e abaixo da média em relação a cada uma das medidas de capital humano. Seria interessante complementar este teste com testes estatísticos adicionais, como o teste de Wilcoxon, não sendo isto aplicado no presente estudo.

As estimações feitas mostram a evolução da sobrevivência das empresas de acordo com determinadas características, não sendo qualquer causa atribuída à variação da mesma. Para

esse efeito, são usados modelos semi-paramétricos e paramétricos que se apresentam de seguida.

4.3.2. Modelos semi-paramétricos e paramétricos

Interessa implementar modelos que permitam estudar o efeito das covariáveis²⁰ na sobrevivência das empresas de forma a complementar a análise anterior. Serão então abordados nesta secção os determinantes da sobrevivência empresarial recorrendo à estimação semi-paramétrica.

O modelo de riscos proporcionais de Cox é o modelo semi-paramétrico mais frequentemente utilizado, contudo, para viabilizar a sua aplicabilidade é necessário que o requisito de proporcionalidade de riscos seja cumprido. O teste ao cumprimento desta proporcionalidade pode ser feito de várias formas mas dado que o uso de métodos gráficos é muito subjetivo (StataCorp, 2011: 157), optou-se pelo teste de Schoenfeld cujas hipóteses são:

$$\left\{ \begin{array}{l} H_0: \text{Os riscos são proporcionais} \\ H_1: \text{Os riscos não são proporcionais} \end{array} \right.$$

Apesar de os modelos paramétricos poderem não fornecer a flexibilidade necessária para representar adequadamente a função de risco (Royston e Lambert, 2011: 14) - fator que conduz à popularidade do modelo de Cox – estes não deixaram também de ser aplicados.

Na Tabela 6 são sumariados os resultados da regressão do modelo de Cox e do modelo paramétrico escolhido - regressões essas que permitem inferir qual a relação entre o capital humano dos fundadores e a probabilidade de sobrevivência das empresas - bem como do respetivo teste de Schoenfeld.

O modelo (1) é um modelo de riscos proporcionais de Cox cuja possibilidade de aplicação é comprovada pelo resultado do teste de Schoenfeld apresentado na Tabela 6. Neste modelo são incluídas variáveis relacionadas com o capital humano do fundador, nomeadamente, o ICH, a experiência no setor no qual a empresa é fundada e a idade. Como variáveis de controlo foram incluídas a dimensão da empresa, o crescimento do setor, a concentração industrial e dummies representativas das regiões Nuts II.

²⁰ A tabela de correlação entre as variáveis pode ser vista no Quadro A.6 em anexo.

O modelo (2) não é mais que a replicação do modelo anterior ao qual acresce a inclusão da variável logaritmizada que mede o crescimento das vendas das empresas, agora aplicados a uma distribuição paramétrica. A não inclusão da variável *cr_vendas* no modelo de Cox prende-se com o facto de esta variável violar o princípio da proporcionalidade dos riscos (ver anexo A.7). Uma alternativa passaria pela aplicação do modelo de Cox estratificado²¹, cuja estratificação é feita pela variável que viola a proporcionalidade. Tal não é possível aplicar diretamente à variável *cr_vendas* porque esta não é uma variável categórica. Desta forma, a estratificação poderia ser implementada dividindo a variável em duas partes, tendo em conta a mediana, não tendo esta solução sido implementada no presente estudo.

A escolha da distribuição de Gompertz foi feita através da técnica de Cox-Snell que nos mostra, entre as várias distribuições paramétricas, qual a mais adequada para estudar a relação entre determinadas variáveis (ver figura A.8.1 em anexo).

Tabela 6. Estimações semi-paramétricas e paramétricas aplicadas aos fundadores.

		Cox (1)	Gompertz (2)
Capital Humano	ICH_fundador	-0.109** (0.054)	-0.119** (0.055)
	expset_fundador	0.012 (0.196)	0.042 (0.194)
	idade_fundador	0.005 (0.005)	0.006 (0.234)
Empresa	dimensao	-0.441 (0.293)	-0.537* (0.302)
	cr_vendas		0.861** (0.035)
Ambiente	cr_empresas	-0.057 (0.057)	-0.073 (0.057)
	IHH	-0.126 (0.078)	-0.121 (0.078)
<i>N</i>		1120	1120
<i>Dummies NUTs II</i>		Sim	Sim
<i>Log Likelihood</i>		-3457.492	-1033.530
<i>AIC</i>		6938.983	2097.060
<i>BIC</i>		6999.236	2172.376
<i>Schoenfeld</i>		0.861	

Notas: Erros padrão entre parêntesis. *Estatisticamente significativo ao nível de 10%, **Estatisticamente significativo ao nível de 5%, ***Estatisticamente significativo ao nível de 1%.

Fonte: Cálculos próprios baseados nos Quadros de Pessoal, MEE.

²¹ Neste modelo assume-se que a amostra se divide em estratos e que o risco base é específico a cada estrato.

Os resultados do modelo de Cox são aqui apresentados de forma a ilustrar como determinadas covariáveis diminuem ou aumentam a taxa de sobrevivência das empresas (Royston e Lambert, 2011).

No modelo (1) interessa analisar o impacto do capital humano dos fundadores sobre a sobrevivência das empresas e verifica-se que o impacto da principal variável – *ICH_fundador* – é relevante. Dado que a interpretação do valor estimado representa a alteração da variável dependente em função de uma alteração unitária em cada uma das variáveis independentes, é visível que um aumento unitário no ICH do fundador leva a uma diminuição de cerca de 10% no rácio de risco ($\exp(-0.109)=0.897$). Uma vez que este índice contempla a educação e a experiência, este resultado vai ao encontro dos resultados de Saridakis, Mole e Storey (2008) de acordo com os quais a educação dos fundadores é uma forma de promoção da sobrevivência das empresas. Esta variável é significativa ao nível de 5% o que contraria os resultados de Gimmon e Levie (2010) que comprovaram que o efeito da educação sobre a sobrevivência das empresas não era significativo. Através deste índice é também alcançada uma concordância com os autores Dahl e Reichstein (2007) e Colombo e Grilli (2005) quanto ao efeito da experiência e ao seu impacto positivo sobre a sobrevivência das empresas. Relativamente ao efeito da experiência do fundador no setor no qual a empresa é fundada pode-se concluir que existe uma relação negativa mas não significativa entre esta variável e a sobrevivência da empresa. Contudo, interessa lembrar que o ICH já inclui uma medida de experiência e mostra que o seu efeito é significativo. Por último, no caso da idade do fundador verifica-se que, à semelhança dos resultados de Thompson (2007), o seu efeito sobre a sobrevivência da empresa não é significativo. Porém, esta componente é também indiretamente incluída no ICH na medida em que este é calculado através da educação, experiência e efeito específico não esquecendo que a experiência foi calculada como uma diferença entre a idade e a educação.

Na distribuição Gompertz, modelo (2), a variável *idade_fundador* continua a não ter um efeito significativo. Sendo o coeficiente da variável *ICH_fundador* de -0.119 verifica-se que uma unidade adicional no ICH leva a uma diminuição de, aproximadamente, 11% ($\exp(-0.119)=0.888$) no risco. A variável de controlo incluída que mostra maior interesse é a dimensão da empresa verificando-se que um aumento unitário do número de estabelecimentos leva a uma diminuição de cerca de 58% ($\exp(-0.537)=0.416$) no rácio de risco aumentando assim a sobrevivência das empresas.

Os resultados das regressões aplicadas aos gestores são apresentados na Tabela 7. Os modelos (3) e (4) incluem as variáveis que medem o ICH dos gestores e a sua experiência quer em cargos de gestão quer no setor no qual a empresa é fundada controlando para as mesmas variáveis que foram utilizadas no caso dos fundadores. O modelo (3) corresponde ao modelo de Cox e o (4), por sua vez, respeita à distribuição Weibull (ver figura A.8.2 em anexo).

Tabela 7. Estimações semi-paramétricas e paramétricas aplicadas aos gestores.

		Cox (3)	Weibull (4)
Capital Humano	ICH_gestor	-0.047 (0.052)	-0.058 (0.053)
	expset_gestor	-0.014 (0.016)	-0.017 (0.164)
	expcergo_gestor	-0.009 (0.040)	-0.024 (0.040)
Empresa	dimensao	-0.490*** (0.296)	-0.597 (0.305)
	cr_vendas		0.092*** (0.036)
	cr_empresas	-0.063 (0.056)	0.077 (0.569)
Ambiente	IHH	-0.146*** (0.077)	-0.142* (0.078)
<i>N</i>		1120	1120
<i>Dummies NUTs II</i>		Sim	Sim
<i>Log Likelihood</i>		-3458.75	-1028.439
<i>AIC</i>		6941.508	2086.878
<i>BIC</i>		7001.761	2162.195
<i>Schoenfeld</i>		0.872	—

Notas: Erros padrão entre parêntesis. *Estatisticamente significativo ao nível de 10%, **Estatisticamente significativo ao nível de 5%, ***Estatisticamente significativo ao nível de 1%.

Fonte: Cálculos próprios baseados nos Quadros de Pessoal, MEE.

No modelo (3), apesar de se verificarem as relações esperadas entre as variáveis de capital humano dos gestores e a sobrevivência das empresas, os seus efeitos não são significativos. Estes resultados induzem a conclusão de que o índice de capital humano dos gestores não tem influência sobre a sobrevivência empresarial. A adequabilidade do modelo (3) é comprovada pelo teste de Schoenfeld cujo resultado não permite rejeitar a hipótese nula a um nível de significância de 5%.

Os resultados do modelo (4), que inclui as mesmas variáveis relacionadas com o capital humano mas ao qual acresce a variável de controlo que respeita ao crescimento das empresas medido através das vendas, mostram impactos positivos ligeiramente mais acentuados das

medidas de capital humano dos gestores sobre a sobrevivência empresarial embora estes permaneçam não significativos. Estes resultados mostram que a escolha da equipa de gestão pode contemplar critérios de seleção baseados na educação e experiência dos indivíduos mas o seu efeito terá um impacto não significativo no objetivo de alcançar uma maior duração por parte da empresa.

Por último foram testados os efeitos do capital humano dos trabalhadores que não são gestores nem fundadores cujos resultados são apresentados na Tabela 8. O modelo (5) é um modelo de riscos proporcionais de Cox e inclui as medidas do ICH dos trabalhadores bem como a sua experiência no setor no qual a empresa se insere. As variáveis de controlo seleccionadas voltam a ser a dimensão e crescimento da empresa, a concentração do setor e a região na qual a empresa se insere. As mesmas variáveis foram aplicadas a uma distribuição paramétrica Gamma, modelo (6), cuja escolha é justificada na figura A.8.3 em anexo tendo sido acrescentado o crescimento da empresa.

Tabela 8. Estimações semi-paramétricas e paramétricas aplicadas aos trabalhadores.

		<i>Cox</i> (5)	<i>Gompertz</i> (6)
Capital Humano	ICH_trabalhador	-0.194*	-0.230**
		(0.109)	(0.105)
	expset_trabalhador	0.074***	0.058***
		(0.023)	(0.021)
Empresa	dimensao	0.306	-0.499*
		(0.328)	(0.301)
	cr_vendas		0.099***
			(0.036)
Ambiente	cr_empresas	0.054	-0.039
		(0.062)	(0.059)
	IHH	0.068	-0.086***
		(0.084)	(0.079)
<i>N</i>		1120	1120
<i>Dummies NUTs II</i>		Sim	Sim
<i>Log Likelihood</i>		-3058.330	-1030.278
<i>AIC</i>		6138.655	2088.555
<i>BIC</i>		6193.887	2158.850
<i>Schoenfeld</i>		0.528	—

Notas: Erros padrão entre parêntesis. *Estatisticamente significativo ao nível de 10%, **Estatisticamente significativo ao nível de 5%, ***Estatisticamente significativo ao nível de 1%.

Fonte: Cálculos próprios baseados nos Quadros de Pessoal, MEE.

Como se pode ver pelos resultados do teste de Schoenfeld aplicado ao modelo (5), a proporcionalidade dos riscos não é, uma vez mais, violada. Pelos resultados deste modelo é visível que um aumento no ICH leva a uma diminuição do rácio de risco em quase 24% (exp(-

0.269)=0.764), o que conduz ao aumento da sobrevivência. Por sua vez, a experiência no setor não apresenta o resultado esperado sendo que um ano adicional nesta medida de capital humano leva a um aumento do rácio de risco de, aproximadamente, 8% ($\exp(0.074)=1.077$).

No modelo (6), sendo este um modelo paramétrico, a variável que viola o princípio da proporcionalidade – *cr_vendas* - pode ser incluída. Os resultados desta distribuição de Gompertz são, em parte, concordantes com o esperado. Relativamente ao ICH dos trabalhadores que não são gestores nem fundadores, este têm um efeito positivo sobre a duração da empresa dado que um aumento do ICH leva a um decréscimo de cerca de 21% ($\exp(-0.230)=0.795$) no rácio de risco, sendo este efeito significativo ao nível de 5%. Por outro lado, a experiência dos trabalhadores no setor mostra um impacto positivo mas não significativo ao contrário do que era esperado. Apesar da não significância da segunda variável, o objetivo de comprovar que a medida criada para medir o capital humano através da educação, experiência e do efeito específico de cada indivíduo é cumprido sendo o seu efeito o esperado independentemente do cargo que os indivíduos ocupam na empresa.

4.4. **Conclusão**

Dado que o objetivo principal do presente estudo é identificar qual o impacto do capital humano sobre a sobrevivência das empresas interessava aplicar um modelo adequado a essa análise. Os vários modelos possíveis foram apresentados neste capítulo e após serem definidos os mais adequados precedeu-se à sua aplicação. Para tal foram incluídas variáveis relacionadas com as empresas, com a força de trabalho, com o ambiente e com o setor no qual as empresas se inserem.

O evento de interesse corresponde à morte das empresas pelo que a variável principal é a duração das mesmas. Uma vez que o evento inicial é dado pelo ano de fundação é eliminada a censura à esquerda e relativamente à censura à direita esta está presente na medida em que no final do período existiam empresas que não tinham ainda morrido. Nestes casos, não é conhecido o tempo de sobrevivência/a duração exata.

Com a aplicação dos vários modelos foram confirmados os efeitos expectáveis para as principais variáveis. O ICH mostrou-se interessante e significativo para os indivíduos que não são gestores tendo um forte impacto sobre a sobrevivência. No que concerne à experiência em cargos similares por parte dos gestores, o seu efeito não se mostrou significativo não sendo este um resultado preocupante na medida em que a experiência geral é incluída no ICH e apresenta o efeito esperado. No que respeita às variáveis que medem a experiência no setor verificou-se

que esta só é significativa no caso dos trabalhadores que não são gestores nem fundadores não sendo o seu impacto coincidente com o previsto.

Por último, relativamente à idade do fundador conjecturava-se que o seu efeito não fosse significativo tendo tal sido confirmado.

Desta forma, é possível responder à questão central deste estudo confirmando que o capital humano da força de trabalho tem um impacto sobre a sobrevivência das empresas.

CAPÍTULO V - CONCLUSÕES

5.1. Síntese

Tal como referem Geroski, Mata e Portugal (2010), a fundação de uma empresa deve ser cuidadosamente pensada e preparada a fim de aumentar a probabilidade de sobrevivência da mesma. O mote da presente dissertação prende-se com a necessidade de discutir a aposta em capital humano como dinamizador da sobrevivência empresarial procurando dar resposta à questão: *Qual o impacto do capital humano na sobrevivência das empresas?* Desta forma, este trabalho propôs-se a contribuir para o estudo da sobrevivência das empresas através de um estudo empírico que recai sobre as empresas fundadas em 1999 analisando a relação entre o capital humano e a sobrevivência das mesmas. Para tal, foram selecionadas algumas empresas portuguesas como unidade de análise e a metodologia foi construída em torno das hipóteses definidas no decorrer da revisão de literatura tendo-se analisado as relações causais entre, por um lado, a sobrevivência/duração das empresas e, por outro, as características do empreendedor, da equipa de gestão e dos restantes trabalhadores. Os Quadros de Pessoal foram a base de dados utilizada pelo facto de contemplarem informação relativa a empresas e trabalhadores e de permitir a associação entre os mesmos. A metodologia desenvolveu-se em vários pontos sendo o primeiro deles a seleção da amostra. Para tal foram identificadas as empresas que respeitam determinados critérios como a não ausência da base de dados por mais que um ano e o número máximo de mortes por empresas igual a um. Empresas que não respeitam os critérios definidos não foram utilizadas neste estudo. Seguiu-se a criação das variáveis necessárias à implementação da análise de duração, variáveis ligadas às características da força de trabalho, das empresas e do ambiente no qual se insere. Por último procedeu-se à implementação de métodos não paramétricos, semi-paramétricos e paramétricos bem como dos testes à sua adequabilidade.

Este capítulo corresponde ao encerramento da dissertação, sendo apresentadas, na secção 5.2., as principais conclusões e, por último, na secção 5.3., são apontadas algumas limitações encontradas e propostas para investigação futura.

5.2. Principais conclusões e recomendações

Vários estudos, entre os quais Dahl e Reichstein (2007), apontam o capital humano dos fundadores e da equipa de gestão como determinantes da sobrevivência empresarial. Contudo, é

visível uma relativa tendência para a não inclusão do impacto do capital humano dos trabalhadores que não sejam fundadores ou gestores. Esta análise não foi descorada neste estudo tendo sido testadas várias hipóteses relacionadas com o capital humano de todos os indivíduos pertencentes às empresas.

A primeira hipótese visava confirmar a existência de uma relação positiva entre a educação e a experiência dos fundadores por um lado, e a sobrevivência das empresas por outro. Para tal foram usadas variáveis que medem o ICH dos fundadores, a sua experiência no setor no qual a empresa é fundada bem como a sua idade. Tendo sido aplicados modelos semi-paramétricos e paramétricos concluiu-se que, tal como era esperado, quanto maior o ICH dos fundadores maior é a sobrevivência das empresas. Pela aplicação dos modelos não existe concordância com a literatura analisada quanto ao impacto da experiência dos fundadores no setor tendo sido verificado um efeito negativo.

No que concerne à segunda hipótese importava testar se a relação existente entre a idade dos fundadores e a sobrevivência das empresas era positiva. Independentemente do modelo aplicado, a idade do fundador apresenta um efeito não significativo. Contudo, importa recordar que o ICH inclui uma medida de experiência geral sendo esta calculada através da idade e da educação dos indivíduos. Assim sendo, a idade é incluída indiretamente na análise através duma variável cujo efeito vai ao encontro daquilo que era expectável.

Relativamente à hipótese 3 que estabelece que há uma relação positiva entre a educação e a experiência dos gestores e a sobrevivência das empresas os resultados demonstram que o efeito destas variáveis não é significativo. Por último, foi testado se a educação e a experiência dos trabalhadores que não são gestores nem fundadores estão positivamente relacionadas com a sobrevivência das empresas. Estas variáveis são incluídas no ICH dos trabalhadores sendo este o índice através do qual se obtêm os impactos mais acentuados sendo que num dos modelos aplicados um aumento unitário nesta medida de capital humano leva a uma redução do rácio de risco de 24%. Ainda no que respeita à experiência dos trabalhadores, o seu efeito sobre a sobrevivência empresarial é significativo sendo o seu efeito negativo o que vai contra o que era esperado.

Em suma, à questão *Qual o impacto do capital humano na sobrevivência das empresas?* pode-se responder que depende da forma como o capital humano é medido. Porém, se aplicar o ICH sugerido neste estudo contendo a educação, a experiência e o efeito específico dos indivíduos comprova-se que um aumento do capital humano dos fundadores, gestores ou

restantes trabalhadores incrementa a capacidade de sobrevivência das empresas. Desta forma, pela importância comprovada da educação percebe-se que é essencial a promoção do ensino, objetivando a redução do abandono escolar, e melhorar a qualidade do mesmo. É também relevante a aplicação de métodos de recrutamento que ponderem a educação e a experiência dos indivíduos.

5.3. Limitações e pistas para investigação futura

Com o presente estudo confirmou-se mais uma vez a complexidade existente na relação entre a sobrevivência e o capital humano.

Uma das dificuldades deste trabalho, à semelhança de outros estudos deste tipo, decorre do facto de não ser possível verificar se a empresa saiu do mercado por ter sido adquirida ou fundida com outra reunindo as condições necessárias para se manter no ativo, ou se saiu porque não conseguiu sobreviver. Desta forma, todas as saídas de empresas foram consideradas como mortes e tentou-se perceber que características das empresas levaram a esse desfecho.

Na construção do ICH é introduzida a experiência geral sendo, para tal, utilizada uma *proxy* da mesma que assume que a escolaridade começa aos 6 anos de idade e que os indivíduos trabalham sem interrupções desde o final dos estudos até à reforma. Esta suposição não tem em conta as situações de desemprego nem as interrupções laborais por parte dos indivíduos do sexo feminino devido a questões familiares, entre outras. Contudo, apesar destas limitações assumiu-se que esta seria a melhor solução perante a informação disponível.

Uma proposta para investigação futura passa pelo foco em setores específicos e na extensão relativamente ao ano de fundação, ou seja, comparação entre empresas fundadas em períodos diferentes. Mais interessante do que esta comparação seria talvez a comparação entre a probabilidade de sobrevivência das empresas incubadas e não incubadas sendo que estas unidades têm como principal objetivo a redução da taxa de mortalidade das empresas apoiadas oferecendo, para tal, facilidades de entrada no mercado bem como de crescimento posterior permitindo às empresas incubadas autossustentarem-se no futuro. Apesar de Medeiros (1998) ter defendido que se as empresas apoiadas por incubadoras aproveitarem o ambiente e os serviços oferecidos têm maior probabilidade de crescimento e de sobrevivência, Schwartz (2010), através dos testes empíricos, concluiu que o facto de uma empresa ter sido apoiada por uma incubadora não aumenta a sua probabilidade de sobrevivência no longo-prazo o que nos poderia induzir a afirmar que as incubadoras não alcançam o principal objetivo identificado pelas

mesmas. Seria então interessante perceber qual das conclusões se aplica à realidade das empresas portuguesas. Na verdade, no início deste trabalho foi feita uma tentativa de inclusão deste determinante tendo sido estabelecido o contacto com um total de 64 unidades de incubação. Porém, dado o reduzido número de respostas – 18 unidades de incubação contribuíram permitindo a reunião das variáveis necessárias²² de 241 empresas incubadas – não foi estudado o impacto deste determinante sobre a sobrevivência das empresas.

A importância da análise do impacto da inovação sobre a sobrevivência das empresas foi também reconhecida por uma vasta lista de autores entre os quais Audretsch (2001), Geroski (1995) e Mata e Portugal (1994). Está assim dado o mote para uma análise que se resume à comparação do impacto do capital humano consoante o nível tecnológico do setor no qual esta se insere. Apesar de na base de dados dos Quadros de Pessoal não estar disponível qualquer informação relativa à inovação em que as empresas incorrem sendo a existência de Diretores do Departamento de I&D a informação mais aproximada, pode ser usada a taxonomia de Pavitt (1984). Esta taxonomia resulta de uma classificação setorial que agrupa as empresas em quatro categorias distintas consoante a origem da tecnologia e a forma de apropriação da mesma, consoante as características dos utilizadores sendo amplamente utilizada como uma medida de intensidade tecnológica. Nos quadros que constam nos anexos A.9 e A.10 são sugeridos os setores que devem constar em cada categoria da referida classificação de forma a ser possível estimar o impacto da inovação sobre a sobrevivência das empresas.

O facto de não ter sido analisado o impacto das unidades de incubação ou da inovação sobre a sobrevivência das empresas não diminui a relevância deste estudo na medida em que são várias as medidas de capital humano incluídas dando-se destaque para a inclusão do ICH que, até à data, não havia sido estudado por nenhum outro autor.

²² Número de Identificação Fiscal da empresa e respetivo período de incubação

BIBLIOGRAFIA

AGARWAL, R. e AUDRETSCH, D. (2001), "Does entry size matter? The impact of the life cycle and technology on firm survival", *The Journal of Industrial Economics*, Vol. 49, pp. 21-43.

ALLISON, P. (1984), "Event History Analysis: Regression for Longitudinal Event Data", Beverley Hills: Sage.

AUDRETSCH, D. (1991), "New-firm survival and the technological regime", *Review of Economics and Statistics*, Vol. 73, pp. 441-450.

AUDRETSCH, D. (1995), "Innovation, growth and survival", *International Journal of Industrial Organization*, Vol.13, pp.441-457.

AUDRETSCH, D. e MAHMOOD, T. (1994), "The Rate of Hazard Confronting New Firm Plants in US", *Review of Industrial Organization*, Vol. 9, pp.41-56.

AUDRETSCH, D. e MAHMOOD, T. (1995), "New Firm Survival: new results using a hazard function", *The Review of economics and statistics*, Vol.60, pp.441-450.

BAPTISTA, R., KARAÖZ, M. e MENDONÇA, J. (2007), "Entrepreneurial Backgrounds, Human Capital and Start-up Success", *Jena Economic Research Papers*, 2007-045.

BATES, T. (1990), "Entrepreneur human capital inputs and small business longevity", *Review of Economics and Statistics*, Vol.72, pp.551-559.

BOSMA, N. VANPRAAG, M., THURIK, R., de WIT, G. (2004), "The value of Human and Social Capital Investments for the Business Performance of startups", *Small Business Economics*, Vol.23, pp.227-236.

BRÜDERL, J., PREISENDÖRFER, P. e Ziegler, R. (1992), "Survival chances of newly founded business organizations", *American Sociological Review*, Vol. 57, pp. 227-242.

CARREIRA, C. e TEIXEIRA, P. (2011), "The Shadow of Death: Analyzing the Pre-Exit Productivity of Portuguese Manufacturing Firms", *Small Business Economics*, Vol.36, pp.337-351.

CEFIS E. e MARSILI, O. (2006), "Survivor: The Role of Innovation in Firms' Survival", *Research Policy*, Vol. 35, pp. 626-641.

CEFIS, E. e MARSILI, O. (2005), "A matter of life and death: innovation and firm survival", *Industrial and Corporate Change*, Vol.14, pp.1167-1192.

CLEVES, M., GUTIERREZ, R., GOULD, W. e MARCHENKO, Y. (2010), "An Introduction to Survival Analysis Using Stata", 3ªEd., College Station, TX: Stata Press.

COLOMBO, M. e GRILLI, L. (2005), "Founders' human capital and the growth of new technology-based firms: A competence-based view", *Research Policy*, Vol.34, pp.795-816.

CONRARIA, L., ALEXANDRE, F. e PINHO, M. (2010), "O euro e o crescimento da economia portuguesa: uma análise contrafactual", *NIPE – WP*, 37.

COX, D. e OAKES, D. (1984), "Analysis of Survival Data", 1ª Ed. London: Chapman & Hall.

COX, D. e SNELL, E. (1968), "A general definition of residuals", *Journal of the Royal Statistical Society B*, Vol.30, pp.248-254.

DAHL, M. e REICHSTEIN, T. (2007), "Are you experienced? Prior Experience and the Survival of New Organizations", *Industry and Innovation*, Vol.14, pp. 497-511.

DAVIDSSON, P. e HONIG, B. (2003), "The role of social and human capital among nascent entrepreneurs", *Journal of Business Venturing*, Vol.18, pp.301-331.

DISNEY, R., HASKEL, J. e HEDEN, Y. (2003), "Entry, Exit and Establishment Survival in UK Manufacturing", *Journal of Industrial Economics*, Vol.51, pp.91-112.

DUNNE, P. e HUGHES, A. (1994), "Age, Size, Growth and Survival: UK Companies in the 1980s", *The Journal of Industrial Economics*, Vol. 42, pp.115-139.

DUNNE, T. ROBERTS, M. e SAMUELSON, L. (1989), "The Growth and Failure of U.S. Manufacturing Plants", *Quarterly Journal of Economics*, Vol. 104, pp. 671-698.

ERIKSSON, T. e KUHN, J. (2006), "Firm spin-off in Denmark 1981-2000 – patterns of entry and exit", *International Journal of Industrial Organization*, Vol.24, pp.1021-1040.

EUROSTAT/OCDE (2007), *Manual on Business Demography Statistics*, Luxemburgo, disponível em <http://www.oecd.org/dataoecd/8/8/39974460.pdf>. Acedido a 10 de abril de 2012

EVANS, D. (1987), "The relationship between firm growth, size, and age: estimates for 100 manufacturing industries", *Journal of Industrial Economics*", Vol.36, pp.567-581.

FOX, J. (2002), "Cox Proportional-Hazards Regression for Survival Data", mimeo, <http://cran.r-project.org/doc/contrib/Fox-Companion/appendix-cox-regression.pdf>.

GEROSKI, P., MATA, J. e PORTUGAL, P. (2010), "Founding Conditions and the Survival of New Firms", *Strategic Management Journal*, Vol. 31, pp. 510-529.

GIMMON, E. e LEVIE, J. (2010), "Founder's human capital, external investment, and the survival of new high-technology ventures", *Research Policy*, Vol.39, pp.1214-1226.

HALL, B. (1987), "The relationship between firm size and firm growth in the U.S. Manufacturing Sector", *Journal of Industrial Economics*, Vol.35, pp.583-605.

HELFAAT, C. e PETERAF, M. (2003), "The Dynamic Resource-based View: Capability Lifecycles", *Strategic Management Journal*, Vol.24, pp.997-1010.

HOLMES, P., STONE, I. e BRAIDFORD, P. (2010), "An Analysis of New Survival using a Hazard Function", *Applied Economics*, Vol. 42, pp. 185-195.

HONJO, Y. (2000), "Business failure of new firms: an empirical analysis using a multiplicative hazards model", *International Journal of Industrial Organization*, Vol.18, pp.557-574.

HOSMER, D. e LEMESHOW, S. (1999), "Applied Survival Analysis", New York: John Wiley & Sons.

ILMAKUNNAS, P. e TOPI, J. (1999), "Microeconomic and Macroeconomic Influences on Entry and Exit of Firms", *Review of Industrial Organization*, Vol.15, pp.283-301.

INE (2003), *Classificação Portuguesa das Atividades Económicas Rev. 2.1*, Lisboa, disponível em <http://metaweb.ine.pt/sine/anexos/pdf/ApresentacaoCAERev2.1.pdf>. Acedido a 13 de Julho de 2012.

INE (2007), *Classificação Portuguesa das Atividades Económicas Rev. 3*, Lisboa, disponível em: <http://metaweb.ine.pt/sine/anexos/pdf/ApresentacaoCAERev3.pdf>. Acedido a 15 de Abril de 2012.

JOVANOVIĆ, B. (1982), "Selection and Evolution of Industry", *Econometrica*, Vol.50, pp. 649-670.

KANIOVSKI, S. e PENEDER, M. (2008), "Determinants of firm survival: a duration analysis using the generalized gamma distribution", *Empirica*, Vol.35, pp.41-58.

KIEFER, N. (1988), "Economic Duration Data and Hazard Functions", *Journal of Economic Literature*, Vol. 26, pp.646-679.

LANCASTER, T. (1990), *The econometric analysis of transition data*, Cambridge. University Press.

MANJÓN-ANTOLÍN, M. e ARAUZO-CAROD, J. (2008), "Firm survival: methods and evidence", *Empirica*, Vol.35, pp.1-24.

MATA, J. e PORTUGAL, P. (1994), "Life Duration of New Firms", *The Journal of Industrial Economics*, Vol. 42, pp. 227-245.

MATA, J. e PORTUGAL, P. (2002), "The survival of new domestic and foreign-owned firms", *Strategic Management Journal*, Vol. 2, pp. 323-343.

MATA, J., PORTUGAL, P. e GUIMARÃES, P. (1995), "The survival of new plants: Start-up conditions and post-entry evolution", *International Journal of Industrial Organization*, Vol. 13, pp. 459 – 481.

MEDEIROS, J. (1998), "Incubadoras de empresas: lições da experiência internacional", *Revista de Administração*, Vol.33, pp.5-20.

PAVITT, K. (1984), "Sectoral Patterns of technical change: towards a taxonomy and a theory", Research Policy, Vol. 13, pp. 343-373.

PHILLIPS, B. e KIRCHHOFF, B. (1989), "Formation, Growth and Survival: Small Firm Dynamics in the US Economy", Small Business Economics, Vol.1, pp.65-74.

PIAZZA-GEORGI, B. (2002), "The role of human and social capital in growth: extending our understanding", Cambridge Journal of Economics, Vol.26, pp.461-479.

PORTELA, M. (2001), "Measuring skill: a multidimensional index", Economics Letters, Vol. 72, pp. 27-32.

PREISENDÖRFER, P. e VOSS, T. (1990), "Organizational Mortality of Small Firms: The Effects of Entrepreneurial Age and Human Capital", Organization Studies, Vol.11, pp.107-129.

ROYSTON, P. e LAMBERT, P. (2011), *Flexible Parametric Survival Analysis Using Stata: Beyond the Cox Model*, Stata Press. College Station, Texas.

SARIDAKIS, G., MOLE, K. e STOREY, D. (2008), "New small firm survival in England", Empirica, Vol.35, pp.25-39.

SCHOENFELD, D. (1982), "Partial residuals for the proportional hazards regression models", Biometrika, Vol.69, pp.239-241.

SCHWARTZ, M. (2010), "A control group study of incubators' impact to promote firm survival", IWH Discussion Papers 11, Halle Institute for Economic Research.

SHANE, S. (2000), "Prior Knowledge and the discovery of entrepreneurial opportunities", Organization Science, Vol.11, pp.448-469.

SIEGEL, R., SIEGEL, E., MACMILLAN, I., (1993), "Characteristics distinguishing high growth ventures", Journal of Business Venturing, Vol. 8, pp. 169–180.

STATACORP. (2011), *Stata: Release 12. Statistical Software*, College Station, TX: StataCorp LP.

STEARNS, T., REYNOLDS, N. e WILLIAMS, M. (1995), "New firm survival: industry, strategy and location", *Journal of Business Venturing*, Vol.10, pp.23-42.

STROTMANN, H. (2007), "Entrepreneurial Survival", *Small Business Economics*, Vol.28, pp. 87-104.

THOMPSON, P. (2007), "Founder Quality and Firm Performance: Implications for Local Development Strategies", Vol.40, pp.97-105.

VERBEEK, M. (2000), *A Guide to Modern Econometrics*, 2ªEd Wiley.

WAGNER, J. (1994), "The post-entry performance of new small firms in German manufacturing industries", *Journal of Industrial Economics*, Vol. 42, pp. 141-154.

WOOLDRIDGE, J. (2010), *Econometric Analysis of Cross Section and Panel Data*, 2ªEd., London. The MIT Press.

ANEXOS

Quadro A.1. Descrição das variáveis utilizadas.

<i>Variável</i>	<i>Descrição</i>
ICH_fundador	ICH dos fundadores
ICH_gestor	ICH dos gestores
ICH_trabalhador	ICH dos trabalhadores
expset_fundador	Experiência passado do fundador no setor no qual a empresa é criada
expset_gestor	Experiência passada do gestor no setor no qual a empresa é criada
expset_trabalhador	Experiência do trabalhador no setor no qual a empresa é fundada
expcargo_gestor	Experiência do gestor em cargos similares
idade_fundador	Idade do fundador
dimensao	Número de Estabelecimentos
cr_trabalhadores	Logaritmo do crescimento das empresas através do número de trabalhadores
cr_vendas	Logaritmo do crescimento das empresas através do volume de vendas
localizacao	11 Região Norte
	15 Região do Algarve
	16 Região Centro
	17 Região de Lisboa
	18 Região do Alentejo
	20 Região dos Açores
	30 Região da Madeira
inda1	Recodificação da indústria
IHH	Logaritmo do Índice de Herfindahl-Hirschman aquando da fundação
cr_empresas	Logaritmo do crescimento absoluto do número de empresas por setor
tempo	Duração da Empresa
died	Morte da empresa
status	Variável de censura

Fonte: Do autor.

Quadro A.2. Correspondência entre as variáveis educação e habil.

<i>educação</i>			<i>habil</i>	
	<i>Ano < 2006</i>		<i>Ano ≥ 2006</i>	
0	110	Não sabe ler nem escrever	111	Não sabe ler nem escrever
	120	Sabe ler e escrever sem possuir o 1º ciclo básico	112	Sabe ler e escrever sem possuir o 1º ciclo básico
2	210	1º ciclo ensino básico (ensino primário 4ª classe)	211	1º ciclo ensino básico (ensino primário 4ª classe)
	220	1º ciclo ensino básico com cursos de índole profissional	212	1º ciclo ensino básico com cursos de índole profissional (ex: cursos da Casa Pia de Lisboa e ex-cursos do Instituto Nacional de Formação Turísticas)
4	310	2º ciclo ensino básico (ensino preparatório, telescola, antigo 2º ano liceu)	221	2º ciclo ensino básico (ensino preparatório, telescola, antigo 2º ano liceu)
	320	2º ciclo ensino básico c/cursos índole profissional	222	2º ciclo ensino básico c/cursos índole profissional (exº: cursos da Casa Pia de Lisboa, cursos complementares de aprendizagem comercial, industrial e agrícola; auxiliares de enfermagem, agentes de educação familiar rural e ex-cursos do Instituto Nacional de Formação Turística)
6	410	3º ciclo ensino básico (antigo 5º ano liceu, 9º ano unificado)	231	3º ciclo ensino básico (antigo 5º ano liceu, 9º ano unificado)
	420	Ensino técnico: c.ger.comercial, c.ger.industrial, c.ger.artes visuais	232	Ensino técnico: c.ger.comercial (ex: curso geral de administração e comércio); Curso geral de artes visuais (artes decorativas)
9	430	3.º ciclo ensino básico c/cursos índole profissional	233	3.º ciclo ensino básico c/cursos índole profissional (ex: ex-cursos do instituto nacional de formação turística)
	440	Cursos das escolas profissionais-nível ii	234	Cursos das escolas profissionais-nível ii (ex: operador agrícola, operador de eletricidade, operador de cerâmica, operador de construção civil e cozinha/pastelaria)

Fonte: Do autor.

Quadro A.2. Correspondência entre as variáveis educ e habil (continuação).

<i>educação</i>		<i>habil</i>	
<i>Ano < 2006</i>		<i>Ano ≥ 2006</i>	
510	Ensino secundário (12.ano) equivalente com cursos de índole profissional, ensino secundário liceu completo	311	Ensino secundário (12.ano) ou equivalente c/cursos de índole profissional (ex: Ex-cursos do Instituto Nacional de Formação Turística); Ensino Secundário Liceal Complementar (11º ano de escolaridade, antigo 7ºano do liceu)
520	Ensino secundário técnico complementar	312	Ensino secundário técnico complementar
530	Ensino secundário técnico profissional	313	Ensino secundário técnico-profissional (cursos profissionais, de especialização e técnico-profissionais ou equivalentes)
540	Cursos das escolas profissionais-nível ii	314	Cursos das escolas profissionais – Nível III (Técnico de contabilidade, técnico de desenho gráfico, técnico de informática, técnico de instalações elétricas, animador socio-cultural)
		414	Ensino pós secundário não superior de nível IV – Formação de professores/formadores e ciências da educação
		421	Ensino pós-secundário não superior de nível IV – Artes
		422	Ensino pós-secundário não superior de nível IV – Humanidades
		431	Ensino pós-secundário não superior de nível IV – Ciências sociais e do comportamento
		432	Ensino pós-secundário não superior de nível IV – Ciências empresariais
		434	Ensino pós-secundário não superior de nível IV – Direito
		438	Ensino pós-secundário não superior de nível IV –Ciências da vida
		442	Ensino pós-secundário não superior de nível IV – Ciências físicas
		444	Ensino pós-secundário não superior de nível IV – Matemática e Estatística
		446	Ensino pós-secundário não superior de nível IV – Informática
		448	Ensino pós-secundário não superior de nível IV – Engenharia e técnicas afins

Fonte: Do autor.

Quadro A.2. Correspondência entre as variáveis educ e habil (continuação).

educação		habil		
Ano < 2006		Ano ≥ 2006		
12		452	Ensino pós-secundário não superior de nível IV – Indústrias transformadoras	
		454	Ensino pós-secundário não superior de nível IV – Arquitetura e construção	
		458	Ensino pós-secundário não superior de nível IV – Agricultura, Silvicultura e Pescas	
		462	Ensino pós-secundário não superior de nível IV – Ciências veterinárias	
		472	Ensino pós-secundário não superior de nível IV – Saúde	
		476	Ensino pós-secundário não superior de nível IV – Serviços sociais	
		481	Ensino pós-secundário não superior de nível IV – Serviços pessoais	
		484	Ensino pós-secundário não superior de nível IV – Serviços de transporte	
		485	Ensino pós-secundário não superior de nível IV – Proteção do ambiente	
		486	Ensino pós-secundário não superior de nível IV – Serviços de segurança	
		499	Ensino pós-secundário não superior de nível IV – desconhecido ou não especificado	
15	614	Bacharelato Formação professores e ciências da educação	514	Bacharelato em Formação professores e ciências da educação
	621	Bacharelato em Artes	521	Bacharelato em Artes
	622	Bacharelato em letras	522	Bacharelato em Humanidades
	631	Bacharelato em Ciências sociais e do comportamento	531	Bacharelato em Ciências sociais e do comportamento
	632	Bacharelato em Jornalismo e informação	532	Bacharelato em Informação e Jornalismo
	634	Bacharelato em comércio e administração	534	Bacharelato em Ciência empresariais
	638	Bacharelato em direito	538	Bacharelato em Direito
	642	Bacharelato em Ciências da vida	542	Bacharelato em Ciências da vida
	644	Bacharelato em Ciências físicas	544	Bacharelato em Ciências físicas
	646	Bacharelato em Matemáticas e estatísticas	546	Bacharelato em Matemáticas e estatísticas
	648	Bacharelato em Ciências informáticas	548	Bacharelato em Informática

Fonte: Do autor.

Quadro A.2. Correspondência entre as variáveis educ e habil (continuação).

<i>educação</i>	<i>Ano < 2006</i>		<i>habil</i>	<i>Ano ≥ 2006</i>	
15	652	Bacharelato em Engenharia e técnicas afins	552	Bacharelato em Engenharia e técnicas afins	
	654	Bacharelato em indústrias de transformação e de tratamento	554	Bacharelato em Indústrias transformadoras	
	658	Bacharelato em Arquitetura e construção	558	Bacharelato em Arquitetura e construção	
	662	Bacharelato em Agricultura, silvicultura e pesca	562	Bacharelato em Agricultura, silvicultura e pesca	
	664	Bacharelato em Ciências veterinárias	564	Bacharelato em Ciências veterinárias	
	672	Bacharelato em saúde	572	Bacharelato em saúde	
	676	Bacharelato em serviços sociais	576	Bacharelato em serviços sociais	
	681	Bacharelato em serviços pessoais	581	Bacharelato em serviços pessoais	
	684	Bacharelato em serviços de transporte	584	Bacharelato em serviços de transporte	
	685	Bacharelato em Proteção do ambiente	585	Bacharelato em Proteção do ambiente	
	686	Bacharelato em Serviços de segurança	586	Bacharelato em Serviços de segurança	
16	699	Bacharelato em Desconhecido ou não especificado	599	Bacharelato em Desconhecido ou não especificado	
	714	Licenciatura em formação de professores e ciências da educação	614	Licenciatura em formação de professores e ciências da educação	
	721	Licenciatura em artes	621	Licenciatura em artes	
	722	Licenciatura em letras	622	Licenciatura em Humanidades	
	731	Licenciatura em ciências sociais e do comportamento	631	Licenciatura em ciências sociais e do comportamento	
	732	Licenciatura em jornalismo e informação	632	Licenciatura em Informação e jornalismo	
	734	Licenciatura em comércio e administração	634	Licenciatura em Ciências Empresariais	
	738	Licenciatura em direito	638	Licenciatura em Direito	
	742	Licenciatura em ciência da vida	642	Licenciatura em ciência da vida	
	744	Licenciatura em ciências físicas	644	Licenciatura em ciências físicas	
	746	Licenciatura em matemáticas e estatísticas	648	Licenciatura em matemática e estatística	
	748	Licenciatura em ciências informáticas	652	Licenciatura em engenharia e técnicas afins	
	752	Licenciatura em engenharia e técnicas afins	654	Licenciatura em Indústrias Transformadoras	
	754	Licenciatura em indústrias de transformação e de tratamento	658	Licenciatura em Arquitetura e construção	

Fonte: Do autor.

Quadro A.2. Correspondência entre as variáveis educ e habil (continuação).

<i>educação</i>		<i>habil</i>		
	<i>Ano < 2006</i>		<i>Ano ≥ 2006</i>	
	758	Licenciatura em Arquitetura e construção	662	Licenciatura em agricultura, silvicultura e pesca
	762	Licenciatura em agricultura, silvicultura e pesca	664	Licenciatura em Ciências veterinárias
	764	Licenciatura em Ciências veterinárias	672	Licenciatura em saúde
	772	Licenciatura em saúde	676	Licenciatura em serviços sociais
	776	Licenciatura em serviços sociais	681	Licenciatura em serviços pessoais
	781	Licenciatura em serviços pessoais	684	Licenciatura em serviços de transporte
	784	Licenciatura em serviços de transporte	685	Licenciatura em proteção do ambiente
	785	Licenciatura em proteção do ambiente	686	Licenciatura em serviços de segurança
	786	Licenciatura em serviços de segurança	699	Licenciatura em Desconhecida ou não especificada
	799	Licenciatura em Desconhecida ou não especificada	714	Mestrado em Formação de professores/formadores e ciências da educação
			721	Mestrado em Artes
			722	Mestrado em Humanidades
			731	Mestrado em Ciências Sociais e do Comportamento
			732	Mestrado em Informação e Jornalismo
			734	Mestrado em Ciências Empresariais
			738	Mestrado em Direito
			742	Mestrado em Ciências da Vida
			744	Mestrado em Ciências Físicas
			746	Mestrado em Matemática e estatística
			748	Mestrado em Informática
			752	Mestrado em Engenharia e técnicas afins
			754	Mestrado em Indústrias Transformadoras
			758	Mestrado em Arquitetura e construção
			762	Mestrado em Agricultura, silvicultura e pescas
			764	Mestrado em Ciências Veterinárias
			772	Mestrado em Saúde

16

Fonte: Do autor.

Quadro A.2. Correspondência entre as variáveis educ e habil (continuação).

<i>educação</i>	<i>habil</i>	
	<i>Ano < 2006</i>	<i>Ano ≥ 2006</i>
16		776 Mestrado em Serviços Sociais
		781 Mestrado em Serviços Pessoais
		784 Mestrado em Serviços de Transporte
		785 Mestrado em Proteção do Ambiente
		786 Mestrado em Serviços de Segurança
		799 Mestrado em Desconhecido ou não especificado
		814 Doutorado em Formação de professores/formadores e ciências da educação
		821 Doutorado em Artes
		822 Doutorado em Humanidades
		831 Doutorado em Ciências Sociais e do Comportamento
		832 Doutorado em Informação e Jornalismo
		834 Doutorado em Ciências Empresariais
		838 Doutorado em Direito
		842 Doutorado em Ciências da Vida
		844 Doutorado em Ciências Físicas
		846 Doutorado em Matemática e estatística
		848 Doutorado em Informática
		852 Doutorado em Engenharia e técnicas afins
		854 Doutorado em Indústrias Transformadoras
		858 Doutorado em Arquitetura e construção
		862 Doutorado em Agricultura, silvicultura e pescas
		864 Doutorado em Ciências Veterinárias
		872 Doutorado em Saúde
		876 Doutorado em Serviços Sociais
		881 Doutorado em Serviços Pessoais
		884 Doutorado em Serviços de Transporte

Fonte: Do autor.

Quadro A.2. Correspondência entre as variáveis educ e habil (continuação).

<i>educação</i>			<i>habil</i>	
	<i>Ano < 2006</i>		<i>Ano ≥ 2006</i>	
16			885	Doutoramento em Proteção do Ambiente
			886	Doutoramento em Serviços de Segurança
			899	Doutoramento em Desconhecido ou não especificado
.	999	Ignorada	999	Ignorada

Fonte: Do autor.

Anexo A.3. Teste de Hausman.

O teste de Hausman foi implementado para comprovar a adequabilidade do modelo de efeitos fixos à estimação do efeito específico dos fundadores, gestores e restantes trabalhadores. A construção do teste resume-se a uma comparação entre os estimadores produzidos pelo modelo de efeitos fixos e pelo modelo de efeitos aleatórios para determinar qual o mais adequado.

Quadro A.3.1. Resultados do teste de Hausman.

	<i>P-valor</i>
Regressão aplicada aos fundadores	0.000
Regressão aplicada aos gestores	0.000
Regressão aplicada aos restantes trabalhadores	0.000

Fonte: Cálculos próprios baseados nos Quadros de Pessoal, MEE.

Dado que o *p-valor* é 0.000 a hipótese nula de que o modelo de efeitos aleatórios fornece estimadores consistentes é rejeitada, ou seja, os efeitos não observáveis são correlacionados com as variáveis explicativas o que faz com que o modelo de efeitos fixos seja o modelo apropriado.

Em suma, os efeitos específicos dos fundadores, gestores e restantes trabalhadores necessários ao cálculo dos respetivos índices de capital humano, podem ser obtidos através do modelo de efeitos fixos.

Quadro A.4. Profissões identificadas como correspondentes a cargos de gestão.

<i>CNP85 – 5 dígitos</i> <i>(Até 1994 – inclusive)</i>		<i>CNP94 – 6 dígitos</i> <i>(A partir de 1994)</i>	
20210	Diretor geral – administração pública Presidente – administração pública Inspetor-geral – administração pública Secretário-geral – administração pública Subdiretor-geral – administração pública	114390	Outros quadros dirigentes, quadros superiores hum. e outr.org.especiais
20310	Diretor de serviços – administração pública	121005	Diretor geral
20320	Chefe de divisão – administração pública	121090	Outros diretores gerais
20390	Outros diretores de serviços e chefes de divisão – administração pública	122105	Diretor de exploração agrícola
21110	Administrador Diretor geral	122110	Diretor de empresa de pescas
21210	Diretor de produção – exceto de empresas agrícolas	122190	Outros diretores de produção e exploração agrícola e similar
21220	Chefe de departamento da produção (chefe de divisão da produção)	122305	Diretor de construção civil e obras públicas
21290	Outros diretores da produção – exceto de empresas agrícolas	122390	Outros diretores de construção e obras públicas
21905	Diretor comercial – exceto comércio por grosso e a retalho	122405	Diretor de comércio retalhista
21910	Diretor financeiro	122410	Diretor de comércio grossista
21915	Diretor de pessoal	122490	Outros diretores comércio grossista e retalhista
21920	Diretor administrativo (diretor de serviços administrativos)	122505	Diretor de hotel
21925	Diretor de informática	122510	Diretor de restaurante
21930	Diretor de aprovisionamento	122590	Outros diretores de restauração e hotelaria
21935	Diretor de exploração – correios e telecomunicações	122705	Diretor de empresa de mediação e serviços
21940	Diretor – estabelecimentos bancários, de seguros ou de operações sobre imóveis Diretor – bancos - serviço de títulos Diretor – bancos – serviço de estrangeiro Diretor – seguros – ramo vida Diretor – seguros – ramo automóvel	122790	Outros diretores empresas mediação e prestação de serviços
21945	Chefe de departamento administrativo (chefe de divisão administrativa, chefe de serviço(s), administrativo(s), chefe de escritório)	122805	Diretor de empresa de serviços.pes.,de limpeza e similares
21990	Outros diretores e quadros dirigentes não classificados em outra parte	122890	Outros diretores de empresa de serviços. pes., limpeza e similares

Fonte: Do autor.

Quadro A.4. Profissões identificadas como correspondentes a cargos de gestão (continuação).

<i>CNP85 – 5 dígitos</i> <i>(Até 1994 – inclusive)</i>		<i>CNP94 – 6 dígitos</i> <i>(A partir de 1994)</i>	
40010	Diretor (gerente) – comércio por grosso	131000	Diretores e gerentes de pequenas empresas (n.e.)
		131100	Diretores e gerentes de agricultura, silvicultura e pesca (n.e)
40020	Diretor (gerente) – comércio a retalho	131190	Outros diretores e gerentes de agricultura, silvicultura e pesca
40090	Outros diretores e gerentes de comércio, por grosso e a retalho	131200	Diretores e gerentes da produção industrial (n.e)
41090	Outros proprietários – gerentes do comércio (comerciantes) por grosso e a retalho	131205	Diretor e gerente da indústria alimentar
50010	Diretor de hotel	131210	Diretor e gerente da indústria extrativa
50020	Diretor de pensão	131215	Diretor e gerente da produção industrial
50030	Diretor de restaurante	131290	Outros diretores e gerentes produção industrial
50090	Outros diretores de hotéis, cafés e restaurantes	131305	Diretor e gerente-construção civil
51010	Proprietário-gerente de hotel	131390	Outros diretores e gerentes da construção civil
51020	Proprietário-gerente de pensão	131405	Diretor e gerente de comércio grossista
51030	Proprietário-gerente de restaurante	131410	Diretor e gerente-comércio retalhista
	Proprietário-gerente de bar		
	Proprietário-gerente de café		
51090	Outros proprietário-gerentes de hotéis, cafés, restaurantes e estabelecimentos similares	131490	Outros diretores e gerentes de comércio grossista e retalhista
60010	Diretor de exploração agrícola	131505	Diretor e gerente de restaurante
60090	Outros diretores e chefes de exploração agrícolas	131510	Diretor e gerente de pensão ou de residencial
		131590	Outros diretores e gerentes de restauração e hotelaria
		131605	Diretor e gerente de transportes
		131610	Diretor e gerente de telecomunicações
		131690	Outros diretores e gerentes de transportes e telecomunicações
		131705	Diretor e gerente de empresas de mediação e serviços
		131790	Outros diretores e gerentes de empresas de mediação e prestação de serviços
		131805	Diretores e gerentes de empresas de serviços pes.,limpeza e similares
		131890	Outros diretores e gerentes de empresas de serviços pes.,limpeza e similares

Fonte: Do autor.

Quadro A.4. Profissões identificadas como correspondentes a cargos de gestão (continuação).

<i>CNP85 – 5 dígitos</i> <i>(Até 1994 – inclusive)</i>		<i>CNP94 – 6 dígitos</i> <i>(A partir de 1994)</i>	
	131905	Diretores e gerentes de org.sanit.,recr.,p.,cultural e outras	
	131990	Outros diretores e gerentes não classificados em outra parte	

Fonte: Do autor.

Quadro A.5. Correspondência entre as variáveis inda1, caemp e caem5r3.

<i>inda1</i>	<i>caemp - CAE_73_Rev.1 – 6 dígitos (De 1986 a 1992 e 1994)</i>	<i>caemp – CAE_Rev. 2.1 – 5 dígitos (Em 1993 e de 1995 a 2006)</i>	<i>caem5r3 – CAE_Rev. 3 – 5 dígitos (A partir de 2007)</i>
2	Mineração	10 Extração de hulha, lenhite e turfa	05 Extração de hulha e lenhite
		11 Extração de petróleo bruto, gás natural e atividades dos serviços relacionados exceto a prospeção	06 Extração de petróleo bruto e gás natural
		12 Extração de minérios de urânio e de tório	07 Extração e preparação de minérios metálicos
		13 Extração e reparação de minérios metálicos	08 Outras indústrias extrativas
		14 Outras indústrias extrativas	09 Atividades dos serviços relacionados com as indústrias extrativas
3	Alimentos, bebidas e tabaco	15 Indústrias alimentares e das bebidas	10 Indústrias alimentares
		16 Indústria do tabaco	11 Indústria das bebidas
			12 Indústria do tabaco
4	Tecidos e peles	17 Fabricação de têxteis	13 Fabricação de têxteis
		18 Indústria do vestuário; preparação, tingimento e fabricação de artigos de peles com pelo	14 Indústria do vestuário
		19 Curtimento e acabamento de peles sem pelo; Fabricação de artigos de viagem, marroquinaria, artigos de correio, seleiro e calçado	15 Indústria do couro e dos produtos de couro
5	Cortiça, papel e madeira à exceção dos móveis.	20 Indústria da madeira e da cortiça e suas obras, exceto mobiliário, fabricação de obras de cestaria e de espartaria	16 Indústrias da madeira e da cortiça e suas obras, exceto mobiliário; fabricação de obras de cestaria e de espartaria

Fonte: Do autor a partir de INE.

Quadro A.5. Correspondência entre as variáveis inda1, caemp e caem5r3 (continuação).

<i>inda1</i>		<i>caemp - CAE_73_Rev.1 – 6 dígitos</i> <i>(De 1986 a 1992 e 1994)</i>		<i>caemp – CAE_Rev. 2.1 – 5 dígitos</i> <i>(Em 1993 e de 1995 a 2006)</i>		<i>caem5r3 – CAE_Rev. 3 – 5 dígitos</i> <i>(A partir de 2007)</i>	
5	<i>Cortiça, papel e madeira à exceção dos móveis.</i>	21	Fabricação de pasta, de papel e cartão e seus artigos	21	Fabricação de pasta, de papel e cartão e seus artigos	17	Fabricação de pasta, de papel, cartão e seus artigos
		22	Edição, impressão e reprodução de suportes de informação gravados	22	Edição, impressão e reprodução de suportes de informação gravados	18	Impressão e reprodução de suportes gravados
						58	Atividades de edição
6	<i>Fabricação de produtos não metálicos</i>	23	Fabricação de coque, produtos petrolíferos refinados e tratamentos de combustível nuclear	23	Fabricação de coque, produtos petrolíferos refinados e tratamentos de combustível nuclear	19	Fabricação de coque, de produtos petrolíferos refinados e de aglomerados de combustíveis
		24	Fabricação de outros produtos químicos	24	Fabricação de produtos químicos	20	Fabricação de produtos químicos e de fibras sintéticas ou artificiais, exceto produtos farmacêuticos
		25	Fabricação de artigos de borracha e de matérias plásticas	25	Fabricação de artigos de borracha e de matérias plásticas	21	Fabricação de produtos farmacêuticos de base e de preparações farmacêuticas
		26	Fabricação de outros produtos minerais não metálicos	26	Fabricação de outros produtos minerais não metálicos	22	Fabricação de artigos de borracha e de matérias plásticas
						23	Fabricação de outros produtos minerais não metálicos
7	<i>Fabricação de produtos metálicos</i>	27	Indústria metalúrgica de base	27	Indústria metalúrgica de base	24	Indústria metalúrgica de base
		28	Fabricação de produtos metálicos, exceto máquinas e equipamentos	28	Fabricação de produtos metálicos, exceto máquinas e equipamentos	25	Fabricação de produtos metálicos, exceto máquinas e equipamentos
		29	Fabricação de máquinas e de equipamentos, n.e.	29	Fabricação de máquinas e de equipamentos, n.e.	26	Fabricação de equipamentos informáticos, equipamento para comunicações e produtos eletrônicos e óticos

Fonte: Do autor a partir de INE.

Quadro A.5. Correspondência entre as variáveis inda1, caemp e caem5r3 (continuação).

<i>inda1</i>	<i>caemp - CAE_73_Rev.1 – 6 dígitos (De 1986 a 1992 e 1994)</i>	<i>caemp – CAE_Rev. 2.1 – 5 dígitos (Em 1993 e de 1995 a 2006)</i>	<i>caem5r3 – CAE_Rev. 3 – 5 dígitos (A partir de 2007)</i>
7 <i>Fabricação de produtos metálicos</i>	30 Fabricação de máquinas de escritório e de equipamento para tratamento automático de informação	30 Fabricação de máquinas de escritório e de equipamento para tratamento de informação	27 Fabricação de equipamento elétrico
	31 Fabricação de máquinas e aparelhos elétricos, n.e.	31 Fabricação de máquinas e aparelhos elétricos, n.e.	28 Fabricação de máquinas e de equipamentos, n.e.
	32 Fabricação de equipamento e aparelhos de radio, televisão e de comunicação	32 Fabricação de equipamento e aparelhos de radio, televisão e de comunicação	29 Fabricação de veículos automóveis reboques, semirreboques e componentes para veículos automóveis
	33 Fabricação de aparelhos e instrumentos médico-cirúrgicos ortopédicos, de precisão de ótica e de relojoaria	33 Fabricação de aparelhos e instrumentos médico-cirúrgicos ortopédicos, de precisão de ótica e de relojoaria	30 Fabricação de outro equipamento de transporte
	34 Fabricação de veículos automóveis, reboques e semirreboques	34 Fabricação de veículos automóveis, reboques e semirreboques	
	35 Fabricação de outro material de transporte	35 Fabricação de outro material de transporte	
	37 Reciclagem	37 Reciclagem	
8 <i>Fabricação de móveis</i>	36 Fabricação de mobiliário; outras indústrias transformadoras, n.e.	36 Indústrias de mobiliário, outras indústrias transformadoras, n.e.	31 Fabricação de mobiliário e de colchões
			32 Outras indústrias transformadoras
10 <i>Fornecimento de gás, eletricidade e água</i>	40 Produção e distribuição de eletricidade, de gás, de vapor e água quente	40 Produção e distribuição de eletricidade, de gás, de vapor e água quente	35 Eletricidades, gás, vapor, água quente e fria e ar frio
	41 Captação, tratamento e distribuição de água	41 Captação, tratamento e distribuição de água	36 Captação, tratamento e distribuição de água

Fonte: Do autor a partir de INE.

Quadro A.5. Correspondência entre as variáveis inda1, caemp e caem5r3 (continuação).

<i>inda1</i>		<i>caemp - CAE_73_Rev.1 – 6 dígitos</i> <i>(De 1986 a 1992 e 1994)</i>		<i>caemp – CAE_Rev. 2.1 – 5 dígitos</i> <i>(Em 1993 e de 1995 a 2006)</i>		<i>caem5r3 – CAE_Rev. 3 – 5 dígitos</i> <i>(A partir de 2007)</i>	
11	<i>Construção</i>	45	Construção	45	Construção	42	Engenharia Civil
						43	Atividades especializadas de construção
12	<i>Comércio retalhista, grossista e reparação de automóveis e motociclos</i>	50	Comércio, manutenção e reparação de veículos, automóveis e motociclos, comércio a retalho de combustíveis para veículos	50	Comércio, manutenção e reparação de veículos, automóveis e motociclos, comércio a retalho de combustíveis para veículos	45	Comércio, manutenção e reparação de veículos automóveis e motociclos
		51	Comércio por grosso e agentes do comércio, exceto de veículos automóveis e de motociclos	51	Comércio por grosso e agentes do comércio, exceto de veículos automóveis e de motociclos	46	Comércio por grosso (inclui agentes), exceto de veículos automóveis e motociclos.
		52	Comércio a retalho exceto de veículos automóveis, motociclos e combustíveis para veículos; reparação de bens pessoais e domésticos	52	Comércio a retalho exceto de veículos automóveis, motociclos e combustíveis para veículos; reparação de bens pessoais e domésticos	47	Comércio a retalho, exceto de veículos automóveis e motociclos
						95	Reparação de computadores e de bens de uso pessoal e doméstico
13	<i>Hotelaria e restauração</i>	55	Alojamento e restauração (restaurantes e similares)	55	Alojamento e restauração (restaurantes e similares)	55	Alojamento
						56	Restauração e similares
14	<i>Transportes, armazenagem e comunicações</i>	60	Transportes terrestres, transportes por oleodutos ou gasodutos (pipelines)	60	Transportes terrestres, transportes por oleodutos ou gasodutos pipelines	49	Transportes terrestres e transportes por oleodutos ou gasodutos
		61	Transportes por água	61	Transportes por água	50	Transportes por água
		62	Transportes aéreos	62	Transportes aéreos	51	Transportes aéreos

Fonte: Do autor a partir de INE.

Quadro A.5. Correspondência entre as variáveis inda1, caemp e caem5r3 (continuação).

<i>inda1</i>		<i>caemp - CAE_73_Rev.1 – 6 dígitos</i> <i>(De 1986 a 1992 e 1994)</i>	<i>caemp – CAE_Rev. 2.1 – 5 dígitos</i> <i>(Em 1993 e de 1995 a 2006)</i>	<i>caem5r3 – CAE_Rev. 3 – 5 dígitos</i> <i>(A partir de 2007)</i>
14	<i>Transportes, armazenagem e comunicações</i>	63 Atividades anexas e auxiliares dos transportes; agências de viagem e de turismo	63 Atividades anexas e auxiliares dos transportes; agentes de viagem e de turismo	52 Armazenagem e atividades auxiliares dos transportes
				79 Agências de viagem, operadores turísticos, outros serviços de reservas e atividades relacionadas
15	<i>Correios e telecomunicações</i>	64 Correios e telecomunicações	64 Correios e telecomunicações	53 Atividades postais e de courier
				61 Telecomunicações
16	<i>Intermediação financeira</i>	65 Intermediação financeira, exceto seguros e fundos de pensões	65 Intermediação financeira, exceto seguros e fundos de pensões	64 Atividades de serviços financeiros, exceto seguros e fundos de pensões
		66 Seguros e fundos de pensões, exceto segurança social obrigatória	66 Seguros, fundos de pensões e de outras	65 Seguros, resseguros e fundos de pensões, exceto segurança social obrigatória
		67 Atividades auxiliares de intermediação financeira	67 Atividades auxiliares de intermediação financeira	66 Atividades auxiliares de serviços financeiros e dos seguros
17	<i>Atividades imobiliárias, alugueres e atividade comercial</i>	70 Atividades imobiliárias	70 Atividades imobiliárias	41 Promoção imobiliária (desenvolvimento de projetos de edifícios); construção de edifícios
		71 Aluguer de máquinas e de equipamentos sem pessoal e de bens pessoais e domésticos	71 Aluguer de máquinas e de equipamentos sem pessoal e de bens pessoais e domésticos	62 Consultoria e programação informática e atividades relacionadas
		72 Atividades informática e conexas	72 Atividades informática e conexas	63 Atividades dos serviços de informação
		73 Investigação e desenvolvimento	73 Investigação e desenvolvimento	68 Atividades imobiliárias

Fonte: Do autor a partir de INE.

Quadro A.5. Correspondência entre as variáveis *inda1*, *caemp* e *caem5r3* (continuação).

<i>inda1</i>	<i>caemp - CAE 73_Rev.1 – 6 dígitos</i> (De 1986 a 1992 e 1994)	<i>caemp – CAE_Rev. 2.1 – 5 dígitos</i> (Em 1993 e de 1995 a 2006)	<i>caem5r3 – CAE_Rev. 3 – 5 dígitos</i> (A partir de 2007)
17 <i>Atividades imobiliárias, alugueres e atividade comercial</i>	74 Outras atividades de serviços prestados principalmente às empresas	74 Outras atividades de serviços prestados principalmente às empresas	69 Atividades jurídicas e de contabilidade
			70 Atividades das sedes sociais e de consultoria para a gestão
			71 Atividades de arquitetura, de engenharia e técnicas afins; atividades de ensaios e de análise técnicas
			72 Atividades de investigação científica e de desenvolvimento
			73 Publicidade, estudos de mercado e sondagens de opinião
			74 Outras atividades de consultoria, científicas, técnicas e similares
			77 Atividades de aluguer
			78 Atividades de emprego
			80 Atividades de investigação e segurança
			81 Atividades relacionadas com edifícios, plantação e manutenção de jardins
			82 Atividades de serviços administrativos e de apoio prestados às empresas
19 <i>Educação</i>	80 Educação	80 Educação	85 Educação
20 <i>Saúde e ação social</i>	85 Saúde e ação social	85 Saúde e ação social	75 Atividades veterinárias
			86 Atividades de saúde humana

Fonte: Do autor a partir de INE.

Quadro A.5. Correspondência entre as variáveis inda1, caemp e caem5r3 (continuação).

<i>inda1</i>		<i>caemp - CAE_73_Rev.1 – 6 dígitos</i> (De 1986 a 1992 e 1994)		<i>caemp – CAE_Rev. 2.1 – 5 dígitos</i> (Em 1993 e de 1995 a 2006)		<i>caem5r3 – CAE_Rev. 3 – 5 dígitos</i> (A partir de 2007)	
20	<i>Saúde e ação social</i>					87	Atividades de apoio social com alojamento
						88	Atividades de apoio social sem alojamento
21	<i>Atividades de serviços comunitários, sociais e pessoais</i>	90	Saneamento, higiene pública e atividades similares	90	Saneamento, higiene pública e atividades similares	37	Recolha, drenagem e tratamento de águas residuais
		91	Atividades associativas diversas, n.e.	91	Atividades associativas diversas, n.e.	38	Recolha, tratamento e eliminação de resíduos; valorização de materiais
		92	Atividades recreativas, culturais e desportivas	92	Atividades recreativas, culturais e desportivas	39	Descontaminação e atividades similares
		93	Outras atividades de serviços	93	Outras atividades de serviços	59	Atividades cinematográficas, de vídeo, de produção de programas de televisão, de gravação de som e de edição de música
						60	Atividades de rádio e de televisão
						90	Atividades de teatro, de música, de dança e outras atividades artísticas e literárias
						91	Atividades das bibliotecas, arquivos, museus e outras atividades culturais
						92	Lotarias e outros jogos de aposta
						93	Atividades desportivas, de diversão e recreativas
						94	Atividades das organizações associativas
						96	Outras atividades de serviços pessoais

Fonte: Do autor a partir de INE.

Quadro A.5. Correspondência entre as variáveis *inda1*, *caemp* e *caem5r3* (continuação).

<i>inda1</i>	<i>caemp - CAE_73_Rev.1 – 6 dígitos</i> (De 1986 a 1992 e 1994)	<i>caemp – CAE_Rev. 2.1 – 5 dígitos</i> (Em 1993 e de 1995 a 2006)	<i>caem5r3 – CAE_Rev. 3 – 5 dígitos</i> (A partir de 2007)
<i>Indústrias eliminadas</i>	01 Agricultura, produção animal, caça e atividades dos serviços relacionados	01 Agricultura, produção animal, caça e atividades dos serviços relacionados	01 Agricultura, produção animal, caça e atividades dos serviços relacionados
	02 Silvicultura, exploração florestal e atividades dos serviços relacionados	02 Silvicultura, exploração florestal e atividades dos serviços relacionados	02 Silvicultura e exploração florestal
	05 Pesca, aquacultura e atividades dos serviços relacionados	05 Pesca, aquacultura e atividades dos serviços relacionados	03 Pesca e aquacultura
	75 Administração pública, defesa e segurança social obrigatória	75 Administração pública, defesa e segurança social obrigatória	97 Atividades das famílias empregadoras de pessoal doméstico
	95 Famílias com empregados domésticos	95 Famílias com empregados domésticos	98 Atividades de produção de bens e serviços pelas famílias para uso próprio
	99 Organismos internacionais e outras instituições extra-territoriais	99 Organismos internacionais e outras instituições extra-territoriais	99 Atividades dos organismos internacionais e outras instituições extraterritoriais

Fonte: Do autor a partir de INE.

Quadro A.6. Coeficientes de correlação entre as variáveis explicativas.

	ICH_fundador	ICH_gestor	ICH_trabalhador	idade_fundador	expcargo_fundador	expset_fundador	expset_gestor	expset_trabalhador	dimensao	regiao	cr_vendas	IHH	cr_empresas
ICH_fundador	1.000												
ICH_gestor	0.781	1.000											
ICH_trabalhador	0.510	0.513	1.000										
idade_fundador	0.481	0.417	0.072	1.000									
expcargo_fundador	0.010	0.039	0.016	0.015	1.000								
expset_fundador	0.012	0.083	-0.015	0.045	0.997	1.000							
expset_gestor	0.027	0.046	0.083	0.153	0.056	0.040	1.000						
expset_trabalhador	0.020	0.045	0.097	0.020	0.067	0.097	0.373	1.000					
dimensao	0.067	0.060	0.064	-0.034	-0.007	-0.008	-0.028	-0.064	1.000				
regiao	0.110	0.132	0.095	0.090	0.018	0.019	-0.022	-0.084	0.062	1.000			
cr_vendas	0.134	0.161	0.392	-0.020	-0.010	-0.018	0.099	0.035	0.046	0.027	1.000		
IHH	-0.129	-0.111	-0.058	-0.057	-0.021	-0.021	-0.015	-0.035	-0.054	0.032	0.062	1.000	
cr_empresas	0.382	0.347	0.392	0.100	0.030	0.030	-0.022	-0.087	0.093	0.085	0.084	-0.601	1.000

Fonte: Cálculos próprios baseados nos Quadros de Pessoal, MEE.

Através da tabela apresentada é possível verificar que as variáveis *expcargo_fundador* e *expset_fundador* são altamente correlacionadas²³ pelo que um modelo que inclua estas duas variáveis pode conduzir a inferências pouco confiáveis. Relativamente às restantes variáveis não se verifica qualquer problema semelhante.

²³ O facto do nível de correlação entre uma variável e ela mesma ser igual a 1 não implica um problema de colinearidade pois é evidente que neste caso haja uma correlação perfeita positiva.

Anexo A.7. Teste da proporcionalidade dos riscos.

Para comprovar a impossibilidade de incluir a variável logaritmizada que mede o crescimento da empresa através das suas vendas, foi aplicado o teste da proporcionalidade dos riscos. Na Tabela A.7.1 são ilustrados os resultados desse teste aplicado a um modelo que inclui as variáveis de capital humano *ICH_fundador*, *expset_fundador* e *idade_fundador* e às variáveis de controlo *dimensao*, *cr_vendas* *IHH* e *cr_empresas*, modelo (1).

Tabela A.7.1. Resultado do teste da proporcionalidade dos riscos.

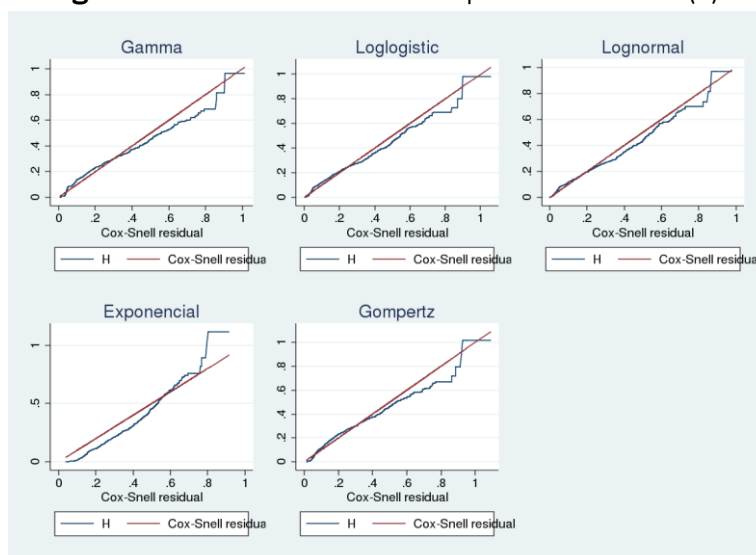
<i>Variável</i>	<i>P-Valor</i>
ICH_fundador	0.966
expset_fundador	0.740
idade_fundador	0.573
dimensao	0.427
cr_vendas	0.000
IHH	0.193
cr_empresas	0.281

Fonte: Cálculos próprios baseados nos Quadros de Pessoal, MEE.

Anexo A.8. Teste de Cox-Snell.

Para decidir qual o modelo mais apropriado entre os vários modelos paramétricos, deve ser feita uma comparação entre os modelos propostos e o estimador de Kaplan-Meier. Esta comparação é apresentada nas figuras que se seguem.

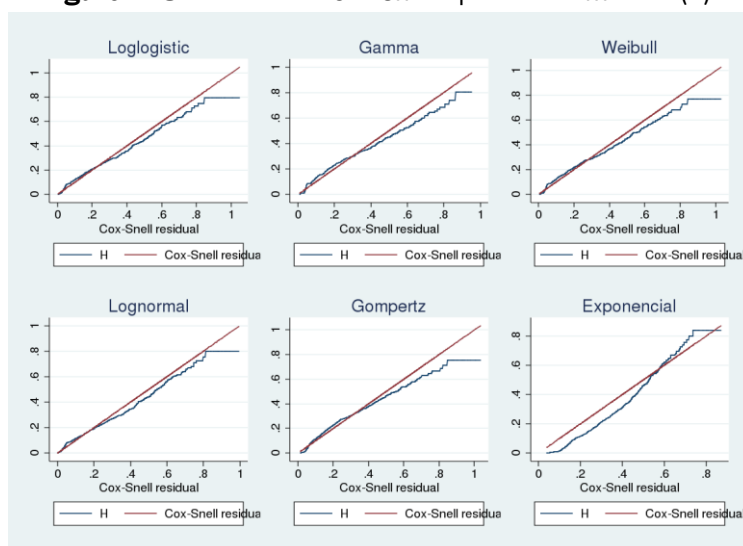
Figura A.8.1. Teste de Cox-Snell aplicado ao modelo (2).



Fonte: Cálculos próprios baseados nos Quadros de Pessoal, MEE.

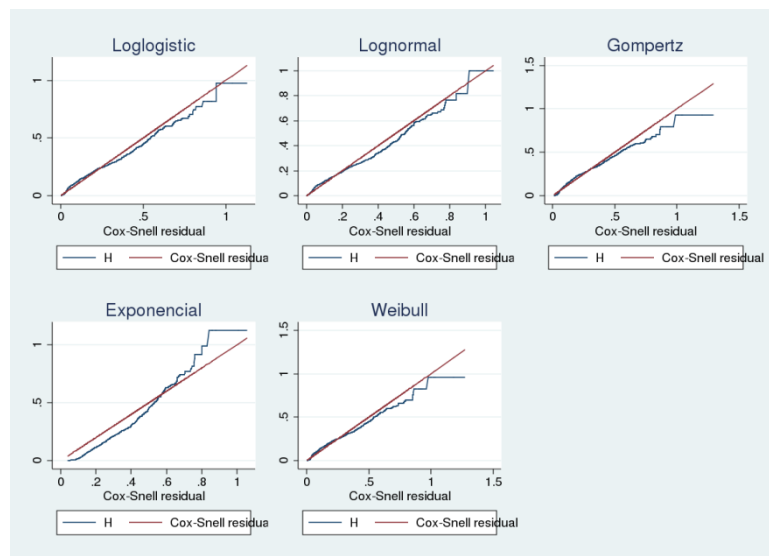
Dado que o modelo que melhor se adequa é aquele com maior aproximação em relação à curva estimada pelo método de Kaplan-Meier, podemos concluir que a distribuição de Gompertz é a mais apropriada.

Figura A.8.2. Teste de Cox-Snell aplicado ao modelo (4).



Fonte: Cálculos próprios baseados nos Quadros de Pessoal, MEE.

Figura A.8.3. Teste de Cox-Snell aplicado ao modelo (6).



Fonte: Cálculos próprios baseados nos Quadros de Pessoal, MEE.

Quadro A.9. Categorias de PAVITT – Indústrias Gerais.

<i>Categoria de PAVITT</i>	<i>Indústria</i>
<i>Dominada pelo fornecedor</i>	Manufatura de têxteis e produtos têxteis; couro e produtos de couro
	Manufatura de madeira e produtos de madeira, celulose, papel e produtos de papel; publicação e impressão
	Construção
	Hotelaria
	Transportes e comunicações
	Administração pública e defesa
<i>Intensiva em escala</i>	Manufatura de produtos alimentares, bebidas e tabaco
	Manufatura de metais básicos
	Manufatura de produtos fabricados com metal, exceto maquinaria e equipamentos
	Manufatura de veículos motorizados, reboques e semirreboques
	Indústrias extrativas
	Produção e distribuição de energia
<i>Fornecedores especializados</i>	Manufatura de maquinaria e equipamentos
<i>Baseada na Ciência</i>	Manufatura de produtos de petróleo e combustível nuclear; produtos químicos e fibras artificiais; produtos de borracha e plástico
	Manufatura de outros minerais não metálicos

Fonte: Do autor.

Quadro A.10. Categorias de PAVITT – Indústrias Específicas.

<i>Categoria de PAVITT</i>	<i>caemp - CAE 73 – 6 dígitos</i>	<i>caemp – CAE_Rev. 2.1 – 5 dígitos</i>	<i>caem5r3 – CAE_Rev. 3 – 5 dígitos</i>
	<i>(De 1986 a 1992 e 1994)</i>	<i>(Em 1993 e de 1995 a 2006)</i>	<i>(A partir de 2007)</i>
<i>Dominada pelo fornecedor</i>	17 Fabricação de têxteis	17 Fabricação de têxteis	13 Fabricação de têxteis
	18 Indústria do vestuário; preparação, tingimento e fabricação de artigos de peles com pelo	18 Indústria do vestuário; preparação, tingimento e fabricação de artigos de peles com pelo	14 Indústria do vestuário
	19 Curtimento e acabamento de peles sem pelo; Fabricação de artigos de viagem, marroquinaria, artigos de correio, seleiro e calçado	19 Curtimento e acabamento de peles sem pelo; Fabricação de artigos de viagem, marroquinaria, artigos de correio, seleiro e calçado	15 Indústria do couro e dos produtos de couro
	20 Indústria da madeira e da cortiça e suas obras, exceto mobiliário, fabricação de obras de cestaria e de espartaria	20 Indústria da Madeira e da cortiça e suas obras, exceto mobiliário, fabricação de obras de cestaria e de espartaria	16 Indústrias da madeira e da cortiça e suas obras, exceto mobiliário; fabricação de obras de cestaria e de espartaria
	21 Fabricação de pasta, de papel e cartão e seus artigos	21 Fabricação de pasta, de papel e cartão e seus artigos	17 Fabricação de pasta, de papel, cartão e seus artigos
	22 Edição, impressão e reprodução de suportes de informação gravados	22 Edição, impressão e reprodução de suportes de informação gravados	18 Impressão e reprodução de suportes gravados
	36 Fabricação de mobiliário; outras indústrias transformadoras, n.e.	36 Indústrias de mobiliário, outras indústrias transformadoras, n.e.	31 Fabricação de mobiliário e de colchões
			32 Outras indústrias transformadoras

Fonte: Do autor.

Quadro A.10. Categorias de PAVITT – Indústrias Específicas (continuação).

<i>Categoria de PAVITT</i>	<i>caemp - CAE 73 – 6 dígitos</i> <i>(De 1986 a 1992 e 1994)</i>		<i>caemp – CAE_Rev. 2.1 – 5 dígitos</i> <i>(Em 1993 e de 1995 a 2006)</i>		<i>caem5r3 – CAE_Rev. 3 – 5 dígitos</i> <i>(A partir de 2007)</i>	
<i>Dominada pelo fornecedor</i>	45	Construção	45	Construção	42	Engenharia Civil
					43	Atividades especializadas de construção
	55	Alojamento e restauração (restaurantes e similares)	55	Alojamento e restauração (restaurantes e similares)	55	Alojamento
					56	Restauração e similares
	60	Transportes terrestres, transportes por oleodutos ou gasodutos (pipelines)	60	Transportes terrestres, transportes por oleodutos ou gasodutos pipelines	49	Transportes terrestres e transportes por oleodutos ou gasodutos
	61	Transportes por água	61	Transportes por água	50	Transportes por água
	62	Transportes aéreos	62	Transportes aéreos	51	Transportes aéreos
	63	Atividades anexas e auxiliares dos transportes; agências de viagem e de turismo	63	Atividades anexas e auxiliares dos transportes; agentes de viagem e de turismo	52	Armazenagem e atividades auxiliares dos transportes
					79	Agências de viagem, operadores turísticos, outros serviços de reservas e atividades relacionadas
	90	Saneamento, higiene pública e atividades similares	90	Saneamento, higiene pública e atividades similares	37	Recolha, drenagem e tratamento de águas residuais
	91	Atividades associativas diversas, n.e.	91	Atividades associativas diversas, n.e.	38	Recolha, tratamento e eliminação de resíduos; valorização de materiais

Fonte: Do autor.

Quadro A.10. Categorias de PAVITT – Indústrias Específicas (continuação).

<i>Categoria de PAVITT</i>		<i>caemp - CAE 73 – 6 dígitos</i>	<i>caemp – CAE_Rev. 2.1 – 5 dígitos</i>	<i>caem5r3 – CAE_Rev. 3 – 5 dígitos</i>
		<i>(De 1986 a 1992 e 1994)</i>	<i>(Em 1993 e de 1995 a 2006)</i>	<i>(A partir de 2007)</i>
<i>Dominada pelo fornecedor</i>	92	Atividades recreativas, culturais e desportivas	92	Atividades recreativas, culturais e desportivas
	93	Outras atividades de serviços	93	Outras atividades de serviços
				39 Descontaminação e atividades similares
<i>Intensiva em escala</i>				59 Atividades cinematográficas, de vídeo, de produção de programas de televisão, de gravação de som e de edição de música
	10	Extração de hulha, lenhite e turfa	10	Extração de hulha, lenhite e turfa
	11	Extração de petróleo bruto, gás natural e atividades dos serviços relacionados exceto a prospeção	11	Extração de petróleo bruto, gás natural e atividades dos serviços relacionados
	12	Extração de minérios de urânio e de tório	12	Extração de minérios de urânio e de tório
	13	Extração e reparação de minérios metálicos	13	Extração e reparação de minérios metálicos
	14	Outras indústrias extrativas	14	Outras indústrias extrativas
	15	Indústrias alimentares e das bebidas	15	Indústrias alimentares e das bebidas
	16	Indústria do tabaco	16	Indústria do tabaco
	27	Indústria metalúrgica base	27	Indústria metalúrgica base
	28	Fabricação de produtos metálicos, exceto máquinas e equipamentos	28	Fabricação de produtos metálicos, exceto máquinas e equipamentos
				05 Extração de hulha e lenhite
				06 Extração de petróleo bruto e gás natural
				07 Extração e preparação de minérios metálicos
				08 Outras indústrias extrativas
				09 Atividades dos serviços relacionados com as indústrias extrativas
				10 Indústrias alimentares
				11 Indústria das bebidas
				12 Indústria do tabaco
				24 Indústrias metalúrgicas de base

Fonte: Do autor.

Quadro A.10. Categorias de PAVITT – Indústrias Específicas (continuação).

<i>Categoria de PAVITT</i>	<i>caemp - CAE 73 – 6 dígitos</i> <i>(De 1986 a 1992 e 1994)</i>	<i>caemp – CAE_Rev. 2.1 – 5 dígitos</i> <i>(Em 1993 e de 1995 a 2006)</i>	<i>caem5r3 – CAE_Rev. 3 – 5 dígitos</i> <i>(A partir de 2007)</i>
<i>Intensiva em escala</i>	34 Fabricação de veículos automóveis, reboques e semirreboques	34 Fabricação de veículos automóveis, reboques e semirreboques	25 Fabricação de produtos metálicos, exceto máquinas equipamentos
	35 Fabricação de outro material de transporte	35 Fabricação de outro material de transporte	29 Fabricação de veículos automóveis reboques, semirreboques e componentes para veículos automóveis
	37 Reciclagem	37 Reciclagem	30 Fabricação de outro equipamento de transporte
	40 Produção e distribuição de eletricidade, de gás, de vapor e água quente	40 Produção e distribuição de eletricidade, de gás, de vapor e água quente	35 Eletricidades, gás, vapor, água quente e fria e ar frio
	41 Captação, tratamento e distribuição de água	41 Captação, tratamento e distribuição de água	36 Captação, tratamento e distribuição de água
	50 Comércio, manutenção e reparação de veículos, automóveis e motocicletas, comércio a retalho de combustíveis para veículos	50 Comércio, manutenção e reparação de veículos, automóveis e motocicletas, comércio a retalho de combustíveis para veículos	45 Comércio, manutenção e reparação de veículos automóveis e motocicletas
	51 Comércio por grosso e agentes do comércio, exceto de veículos automóveis e de motocicletas	51 Comércio por grosso e agentes do comércio, exceto de veículos automóveis e de motocicletas	47 Comércio a retalho, exceto de veículos automóveis e motocicletas
<i>Fornecedores especializados</i>	52 Comércio a retalho exceto de veículos automóveis, motocicletas e combustíveis para veículos; reparação de bens pessoais e domésticos	52 Comércio a retalho exceto de veículos automóveis, motocicletas e combustíveis para veículos, reparação de bens pessoais e domésticos	
	29 Fabricação de máquinas e de equipamentos, n.e.	29 Fabricação de máquinas e de equipamentos, n.e.	26 Fabricação de equipamentos informáticos, equipamento para comunicações e produtos eletrónicos e óticos

Fonte: Do autor.

Quadro A.10. Categorias de PAVITT – Indústrias Específicas (continuação).

<i>Categoria de PAVITT</i>	<i>caemp - CAE 73 – 6 dígitos</i> <i>(De 1986 a 1992 e 1994)</i>	<i>caemp – CAE_Rev. 2.1 – 5 dígitos</i> <i>(Em 1993 e de 1995 a 2006)</i>	<i>caem5r3 – CAE_Rev. 3 – 5 dígitos</i> <i>(A partir de 2007)</i>
<i>Fornecedores especializados</i>	30 Fabricação de máquinas de escritório e de equipamento para tratamento automático de informação	30 Fabricação de máquinas de escritório e de equipamento para tratamento de informação	27 Fabricação de equipamento elétrico
	31 Fabricação de máquinas e aparelhos elétricos, n.e.	31 Fabricação de máquinas e aparelhos elétricos, n.e.	28 Fabricação de máquinas e de equipamentos, n.e.
	32 Fabricação de equipamento e aparelhos de radio, televisão e de comunicação	32 Fabricação de equipamento e aparelhos de radio, televisão e de comunicação	31 Fabricação de mobiliário e de colchões
	33 Fabricação de aparelhos e instrumentos médico-cirúrgicos ortopédicos, de precisão de ótica e de relojoaria	33 Fabricação de aparelhos e instrumentos médico-cirúrgicos ortopédicos, de precisão de ótica e de relojoaria	
	23 Fabricação de coque, produtos petrolíferos refinados e tratamentos de combustível nuclear	23 Fabricação de coque, produtos petrolíferos refinados e tratamentos de combustível nuclear	19 Fabricação de coque, de produtos petrolíferos refinados e de aglomerados de combustíveis
	24 Fabricação de outros produtos químicos	24 Fabricação de produtos químicos	20 Fabricação de produtos químicos e de fibras sintéticas ou artificiais, exceto produtos farmacêuticos
<i>Baseada na Ciência</i>	25 Fabricação de artigos de borracha e de matérias plásticas	25 Fabricação de artigos de borracha e de matérias plásticas	21 Fabricação de produtos farmacêuticos de base e de preparações farmacêuticas
	26 Fabricação de outros produtos minerais não metálicos	26 Fabricação de outros produtos minerais não metálicos	22 Fabricação de artigos de borracha e de matérias plásticas
			23 Fabricação de outros produtos minerais não metálicos

Fonte: Do autor.